أساليب البحث العلمي في العلوم الإداريّة

الدكتور
علي سليم العلاونة
جامعة مؤتة

بكالوريوس : محاسبة / احصاء - جامعة اليرموك
ماجستير : ادارة أعمال - جامعة ألباما - أمريكا
دكتوراة : ادارة انتاج وعمليات - جامعة شيفيلد - انجلترا

الطبعة الأولى
1996 م - 1416 هـ

دار الفكر للنشر والتوزيع - عمّان

بسم الله الرحمن الرحيم

﴿ وَقُلِ ٱعْمَلُوا۟ فَسَيَرَى ٱللَّهُ عَمَلَكُمْ وَرَسُولُهُۥ وَٱلْمُؤْمِنُونَ ﴾

صدق الله العظيم

سورة التوبة/ آية (١٠٥)

أساليب

البحث العلمي

في العلوم الإدارية

الطبعة الأولى

1996 م - 1416 هـ

DAR AL-FIKR

Printing - Publishing - Distributing

دار الفكر

للطباعة والنشر والتوزيع

سوق البتراء (الحجيري) هاتف ٦٢١٩٣٨ - فاكس ٦٥٤٧٦١ ص.ب ١٨٣٥٢٠ عمّان ١١١١٨ الأردن

Tel.: 621938 - Fax.: 654761 - P.O.Box: Amman 11118 Jordan

إهداء

إلى
ذكرى الوالد والوالدة
وإلى
الزوجة والأبناء
وائل ومحمد ورغد وشهد

محتويات الكتاب

تمهيد

إن عملية البحث في وقتنا الحاضر لم تعد كما كانت عليه في الوقت السابق معتمدة على معرفة الشخص وقدراته وما الى ذلك، بل أصبحت طريقة تحكمها الأسس الموضوعية من لحظة البدء بالتفكير بتنفيذ البحث الى مرحلة الانتهاء من كتابة تقرير البحث. ولعل السبب في ذلك هو عظم أو كبر التحدي والتغير الذي يواجهنا في أيامنا هذه والذي جلب معه الحاجة لامتلاك المدراء ومتخذي القرارات المهارات والمقدرة المناسبة حتى يتمكنوا من الاستجابة للضغوطات والظروف التي تواجههم.

وحتى يتمكن المدراء من تحسين قدرتهم على مواجهة التحديات والصعوبات التي تواجههم، فانه لا بد من توفير المعلومات المناسبة والكافية حتى تكون لهم عوناً على تحسين نوعية القرارات الادارية لمواجهة تلك التحديات. ان الوظيفة الأساسية للأبحاث الادارية هي تزويد المعلومات المناسبة عن الموضوع مدار الاهتمام للمدراء أو متخذي القرارات. وعليه، فان المدير الاداري أو متخذ القرار الذي لا يستطيع استخدام أو تقييم الأبحاث الادارية في الوقت الحاضر، كالمدير العام الذي لا يستطيع فهم الورقة الربحية للشركة. بمعنى آخر، ان مقدرة كلا النوعين من الأشخاص في تنفيذ أعمالهم بصورة فعالة تكون محدودة جداً.

ونظراً للأهمية الكبرى للأبحاث الادارية في تزويد المعلومات المناسبة ولشح المكتبة العربية بشكل عام والأردنية بشكل خاص من هذا النوع من المؤلفات، فقد كان هذا حافزاً للمؤلف لبذل الجهد من أجل اخراج هذا الكتاب المتواضع في أساليب البحث العلمي . وأهم ما يتميز به هذا الكتاب هو سهولة القراءة والفهم من خلال العناية الكافية التي أُعطيت لتوضيح وشرح القضايا الأساسية التي يتضمنها الكتاب حتى يتمكن أي قاريء من الاستفادة من محتوياته. لذلك، فان

الأهداف أو الأغراض الأساسية لهذا الكتاب هي ما يلي :

١ ـ تحسين وتوسيع أفق المدراء ومتخذي القرارات على فهم وتحليل الأبحاث الادارية من خلال الشرح المفصل لكل خطوة من خطوات عملية البحث من أجل تزويدهم بالمعلومات الضرورية التي تكفل حل المشكلة أو اتخاذ القرار المناسب بفعالية.

٢ ـ تزويد الطالب الجامعي بكتاب حول أساليب البحث العلمي المختلفة بحيث يتناول شرح العناصر الأساسية بشكل مفصل وبسيط ولا يشترط المعرفة السابقة للاحطاء من أجل فهمه، لتحسين مقدرته البحثية والتحليلية حتى يتمكن من الحكم على الأشياء بموضوعية.

٣ ـ يمكن أن يعتبر هذا الكتاب كمرجع أولي لأولئك الأشخاص الذين يريدون أن يتخصصوا في كيفية عمل الأبحاث الادارية.

وعليه، فقد تم تصميم هذا الكتاب بأربعة عشر فعلاً، والتي كانت كما يلي :

١ ـ مقدمة في البحث العلمي. ٢ ـ تصميم البحث. ٣ ـ المعاينة. ٤ ـ العينة العشوائية البسيطة وحجم العينة. ٥ ـ طرق العينات الأكثر تعقيد. ٦ ـ عملية القياس. ٧ ـ طرق جمع البيانات. ٨ ـ تصميم أشكال جمع البيانات. ٩ ـ اعداد البيانات. ١٠ ـ تحليل البيانات. ١١ ـ اختبار الفرضيات. ١٢ ـ الانحدار البسيط والارتباط. ١٣ ـ الانحدار المتعدد والارتباط. ١٤ ـ كتابة تقرير البحث.

آملاً أن يحقق هذا الكتاب الأغراض التي من أجلها تم اعداده.

والله الموفّق

المؤلف

الفصل الأول

مقدمة في البحث العلمي

Introduction to Scientific Research

مقدمة في البحث العلمي
Introduction to Scientific Research

لقد بدأ الانسان البحث منذ وجوده على هذه البسيطة حتى يتمكن من تلبية حاجاته ورغباته المختلفة المادية منها والمعنوية. ولكن البحث في الماضي لم يكن علمياً ويتمتع بمنهجية واضحة كما هو عليه الآن لبساطة الحاجات والرغبات التي كان يبحث عنها الانسان البدائي والتي كانت تتحقق بذلك النوع من الأبحاث. وكنتيجة للتقدم والتعقيد الذي حصل في حياتنا الحالية سواء في الظروف التقنية أوَ الظروف الاجتماعية أو الاقتصادية أو السياسية كان لا بد ان تتغير معها طريقة البحث بشكل يضمن تحقيق حاجات ورغبات الانسان في هذا العصر. لذلك نجد ان البحث في ايامنا تغيّر كثيراً مقارنة مع البحث البدائي وأصبح يعتمد على طرق واضحة ومنهجية محددة في الوصول الى الحقائق أو في اكتشاف الظواهر. ولأهمية البحث في أيامنا هذه نجده جزءاً لا يتجزأ من حياة أية أمة أو مؤسسة أو هيئة بحيث أصبح يُخصص له موارد مالية وبشرية للاعتماد عليه في اتخاذ القرارات بمختلف أنواعها.

ان المشكلة الرئيسية التي تواجه معظم المؤسسات في أيامنا هذه هو كيفية مراقبة حاجات الانسان ومعرفة المشاكل التي تواجهه في الحاضر والتنبؤ بها في المستقبل وكإستجابة لهذا الوضع والمساعدة في تنبؤ المستقبل فإن مبدأ النظام الرسمي في جمع المعلومات للتسهيل أو المساعدة في عملية اتخاذ القرارات أصبح ضرورة.

ـ تعريف البحث Research Definition

يوجد العديد من التعريفات الجيدة للبحث والتي منها ان البحث عبارة عن عملية منظمة في جمع البيانات وتسجيلها وتحليلها لحل مشكلة ما. ولكن لأغراضنا،

نود أن يتضمن تعريف البحث اربعة عناصر رئيسية. وهذه العناصر هي :

(١) طريقة منظمة Systematic Approach

(٢) الموضوعية Objectivity

(٣) المعلومات Information

(٤) اتخاذ القرار Decision Making

وعليه يمكن تعريف البحث على انه طريقة منظمة وموضوعية في جمع البيانات وتسجيلها وتحليلها لاستخلاص وتطوير المعلومات لتزويدها للمؤسسة أو صاحب القرار لاستخدامها في عملية اتخاذ القرار.

تشير الطريقة المنظمة في هذا التعريف لضرورة تنظيم وتخطيط مشروع البحث من خلال استخدام النواحي الاستراتيجية والفنية في تصميم البحث وكذلك توقع البيانات المُراد تجميعها وأنواع الاختبارات التي يجب أن تستخدم في التحليل. وتتضمن الموضوعية وجوب محاولة عدم التحيز في انجاز البحث. بمعنى آخر استخدام الطريقة العلمية في عملية البحث.

ـأنواع البحث Types of Research

يمكن تقسيم الدارسات البحثية الى قسمين :

١ ـ البحث الأساسي أو النظري والذي يهدف الى توسيع حدود المعرفة بالنسبة لظاهرة معينة أو مشكلة معينة. وهناك بعض التساؤلات حول امكانية استخدام مثل هذا النوع من المعرفة في العملية الادارية.

٢ ـ البحث التطبيقي (التحرّي) والذي يهتم بمساعدة الادارة في اتخاذ قرارات افضل. وتكون هذه الأبحاث موجّهة نحو حالات محددة من المؤسسة ،نتائجها تعمل كدليل في عملية اتخاذ القرارات.

ويمكن التميز ما بين البحث الأساسي والتطبيقي من خلال شمولية البحث. ان من الخصائص المرغوبة في البحث النظري هو الشمولية والكمالية Thorough and Complete . ولكن الشمولية في البحث التطبيقي تكون محددة بالمعلومات المطلوبة لمتخذ القرار.

وحتى نتمكن من اجراء البحث سواء كان نظرياً أم تطبيقياً، فانه لا بد من توفر المال والجهد والوقت والمنهجية. وتعرّف المنهجية على أنها التحري الدقيق وباجراءات منظمة ومنسقة للحصول على نتائج لاستخلاص المعلومات المطلوبة ومن ثم معرفة الحقيقة. وعليه فان الوقت والمال والجهد المطلوب لعمل دراسة شاملة قد لا يكون ضروري مقارنة مع الدراسة التطبيقية اذا ما اخذنا بعين الاعتبار المعلومات المطلوبة لاتخاذ قرار كفيء وفعّال. لذلك ان التركيز في هذا الكتاب سوف يكون بشكل رئيسي على الأبحاث التطبيقية. ولكن هذا لا يمنع من استخدام بعض المبادىء والطرائق البحثية على الأبحاث النظرية.

ـ العلاقة بين الادارة والبحث The Management - Research Relationship

تعتمد فعالية نظام البحث على حجم العلاقة ما بينها وبين الادارة الطالبة للبحث. ويوجد هناك العديد من العوامل التي تؤثر على هذه العلاقة مثل الاختلافات ما بين بعض المدراء والباحثون بالنسبة لبعض المسؤوليات وأهداف العمل والخلفية التعليمية.

ان نظام البحث يتطلب مهارات ومعرفة من الأشخاص الباحثين المتخصصين. لذلك قد يحتوي قسم البحث والتطوير في اي مؤسسة على كفاءات متخصصة في مراحل عملية البحث المختلفة مثل تصميم الاستبيان ومسح مجتمع الدراسة واعداد البيانات وتحليلها واعداد التقارير. ولكن التخصص يتضمن معرفة الشخص الباحث في طبيعة مشكلة البحث قيد الدراسة بالاضافة الى معرفته في الاحصاء والرياضيات وعلم

النفس واستخدام الحاسوب والاقتصاد حتى يصبح باحث متخصص مؤهل. ولكن غالباً ما تكون خبرة الباحثين بالنسبة لدور الأبحاث التطبيقية في عملية اتخاذ القرارات الادارية قليلة.

اذا كان الشخص متخصص في البحث وليس موجّه ادارياً فانه يمكن الاستفادة منه في عملية البحث دون التدخل في عملية تحديد الحاجة للبحث. فالشخص الباحث في هذه الحالة قد يفشل في طرح أسئلة موجّه بالنسبة لحالة القرار وبالتالي قد لا تساهم نتائج بحثه في عملية اتخاذ القرار.

ـ صفات الباحث

يتوجب على الشخص الباحث ان يتميز بخصائص وصفات شخصية تميزه عن غيره حتى يتمكن صاحب القرار من الاستفادة من نتائج الأبحاث التي يقوم بها في عملية اتخاذ القرارات بالنسبة للمشكلة قيد البحث . ويمكن تقسيم هذه الصفات والتي يجب ان يتحلّى بها الشخص الباحث الى قسمين : (١) صفات خلقية. (٢) صفات علمية . وفيما يلي شرح لهاتين الخاصيتين :

١ ـ صفات خلقية.

تتعلّق الصفات الخلقية بتلك الخصائص ذات العلاقة بالباحث كانسان. وأول هذه الصفات التي يجب توفرها في الشخص الباحث هي الصدق والأمانة في تسجيل وتفريغ البيانات المستخدمة في حل مشكلة البحث قيد الدراسة. كما يتوجب على الشخص الباحث ان يحب عمله الذي يقوم به والذي ينعكس من خلال قوة الرغبة عنده في القيام بعملية البحث لأن ذلك ينعكس بشكل مباشر على درجة الموضوعية في نتائج البحث. كما نتوقع ان يكون الباحث موضوعياً في تعامله مع الناس وموضوعياً في مناولة المشكلة ودقيقاً في تسجيل البيانات

اللازمة وذو قدرة قوية على الملاحظة حتى يتمكن الباحث من عكس الواقع من خلال البحث قيد الدراسة.

٢ ـ صفات علمية.

تتعلّق الصفات العلمية بتلك الخصائص ذات العلاقة بالباحث كعالم. ويمكن تقسيم هذه الصفات الى نوعين :

أ ـ الصفات العلمية العامة :

وتتمثل هذه الصفات بالخصائص التي يجب ان تتوفر في كل باحث وبغض النظر عن نوع الأبحاث التي يقوم بها. وتشمل هذه الصفات على ما يلي :

١ ـ مقدرة الباحث على القيام بعملية تصميم البحث بطريقة جيدة حتى يتمكن من : (أ) الحصول على البيانات المناسبة لتحقيق أهداف الدراسة.
(ب) تحديد الاختبارات والتحاليل اللازمة لتحقيق أهداف البحث.
(جـ) اختيار الطريقة الأمثل في جمع البيانات المطلوبة.
(د) تنفيذ اجراءات البحث بطريقة اقتصادية.

٢ ـ مقدرة الشخص الباحث في ربط البيانات التي تمّ تجميعها والمعلومات المطلوبة لعملية اتخاذ القرار.

٣ ـ المقدرة على عمل الأبحاث بموضوعية لتجنب التحيز. بمعنى آخر البعد عن العادات والتقاليد والخبرة الشخصية في عملية البحث حتى يتمكن من الوصول الى نتائج يمكن استخدامها في حل مشكلة البحث قيد الدراسة.

ب ـ صفات علمية خاصة.

تتعلق الصفات العلمية الخاصة بتلك الخصائص التي يجب أن تتوفر بالشخص الباحث وذات علاقة بموضوع البحث قيد الدراسة. ومن هذه الخصائص المعرفة

النظرية بموضوع البحث قيد الدراسة. والمعرفة النظرية مهمة جداً لتشكيل أهداف الدراسة وتحديد المتغيرات ذات العلاقة بتحقيق أهداف الدراسة قيد البحث حتى يتمكن أصحاب القرارات من الأستعانة في نتائج الدراسة على عمل القرارات. كما تتعلق هذه الخصائص بالمعرفة بطرائق البحث المختلفة وكيفية المفاضلة ما بين هذه الطرائق حتى يتم استخدام طريقة البحث الأمثل في حل مشكلة البحث قيد الدراسة. هذا بالاضافة الى مقدرة الباحث في تحديد المجتمع الأمثل لجمع البيانات لضمان جمع بيانات تتفق والمعلومات المطلوبة.

ـ عملية البحث The Research Process

يمكن النظر الى مشروع البحث كسلسلة من الخطوات المتتابعة والتي يطلق عليها اسم عملية البحث والشكل (١ - ١) يوضح الخطوات في هذه العملية. وحتى نتمكن من القيام بالبحث بكفاءة وفعالية، فانه من الضروري توقع خطوات البحث وادراك العلاقات فيما بينها قبل البدء بعملية البحث. وسوف نقوم الآن بتوضيح خطوات عملية البحث بشكل بسيط حتى نبين تسلسل هذه الخطوات ودرجة اعتمادها على بعضها البعض، ثم نقوم بتفصيل هذه الخطوات بالفصول اللاحقة.

شكل (١ ـ ١)

١ ـ تحديد الحاجة للمعلومات

1- Establish the need for Information.

↓

٢ ـ تحديد أهداف البحث والمعلومات المطلوبة

2- Specify Research Objectives and Information Needs.

↓

٣ ـ صياغة الفرضيات

3- Hypotheses Formulation.

↓

٤ ـ تحديد مصادر البيانات

4- Determine Sources of Data.

↓

٥ ـ تطوير أدوات جمع البيانات

5- Develop the Data Collection Forms.

↓

٦ ـ تصميم العينة

6- Design the Sample.

↓

٧ ـ جمع البيانات

7- Data Collection.

↓

٨ ـ اعداد البيانات

8- Process the Data.

↓

٩ ـ تحليل البيانات

9- Analyze the Data.

↓

١٠ ـ عرض نتائج البحث

10- Present Research Results.

١ ـ الحاجة للمعلومات.

ان بيان الحاجة لمعلومات البحث بالتأكيد هي أول خطوة من خطوات عملية البحث. ان الطلب الأوّلي للمدير بالمساعدة يبين الحاجة لمعلومات البحث. وحتى يتمكن الباحث من القيام بالبحث فيتوجب عليه معرفة وفهم وبشكل كامل سبب أو أسباب الحاجة الى البحث. فالمدير هو المسؤول عن توضيح الظروف المحيطة عن الشيء المطلوب له معلومات لأن ذلك سوف يسهّل في عملية اتخاذ القرار Decision- making process واذا كان المطلوب من البحث تزويد معلومات مهمة لعمل القرار، فان ذلك يتطلب تحديد الحاجة للمعلومات البحثية وبدقة. (الفصل الرابع يوضح بالتفصيل).

ان المدراء غالباً ما يواجهوا أعراض المشكلة وليس المشكلة المطلوب تجميع بيانات عنها. لذلك فان تحديد الحاجة للمعلومات البحثية مرحلة حرجة وصعبة في عملية البحث. لهذا فانه من المهم جداً اعادة النظر في هذه المرحلة من مشروع البحث لضمان نتائج بحث تساعد في اتخاذ القرار.

٢ ـ أهداف البحث والمعلومات المطلوبة.

عندما تحدد الحاجة لمعلومات البحث وبوضوح، فان اهتمام الشخص الباحث يجب أن يتحول الى تحديد أهداف للبحث المنوي عمله وتطوير قائمة محددة بالمعلومات المطلوبة. كما يجب على أهداف البحث أن تجيب عن السؤال التالي: «لماذا يجب أن نقوم بالبحث؟». بشكل عام تحدد أهداف البحث كتابةً قبل البدء بتنفيذ البحث. ويجب أن تكون المعلومات المطلوبة قادرة على تحقيق أهداف البحث. ويمكن التفكير بالمعلومات المطلوبة من ناحية عملية كقائمة مفصلة لأهداف الدراسة.

٣ ـ صياغة الفرضيات.

يتوجب على الشخص الباحث بعد تحديد أهداف الدراسة أن يقوم بعملية صياغة الفرضيات التي يتوقع أن تساعد في تفسير مشكلة البحث. والفرضية عبارة عن جملة تقترح من خلالها حل معين للمشكلة قيد الدراسة. لذلك، يتوجب صياغة الفرضية بشكل دقيق جداً نتمكن من القيام بالحكم عن الاقتراح الموجود بالفرضية. وتتخذ الفرضية شكلان عند صياغتها هما صيغة النفي وصيغة الاثبات. صيغة النفي تنفي وجود علاقة أو فرق بين متغيرين أو أكثر. بينما صيغة الاثبات تؤكد وجود علاقة أو فرق بين متغيرين. وغالباً ما تكتب الفرضية بصيغة النفي وتدعى الفرضية الأساسية والتي بها يبدأ اختبار الفرضيات.

٤ ـ مصادر البيانات.

ان الخطوة التي تتبع تحديد اهداف الدراسة وقائمة مفصلة للمعلومات المطلوبة هي تحديد مصادر البيانات، ومعرفة ما اذا كانت تتوفر داخل المؤسسة التي تدعم البحث أم خارجها. فالمصادر الداخلية تتضمن الأبحاث السابقة وملفات وسجلات المؤسسة. بينما تتضمن المصادر الخارجية للبيانات تقارير الأبحاث التجارية ومجلات وتقارير الصناعة والتقارير الحكومية ... الخ. فاذا وجد الشخص الباحث البيانات المناسبة للمعلومات المطلوبة ، فيتوجب عليه فحص تصميم البحث للتاكد من دقتها لأن سمعة المؤسسة التي تجمع وتحلل البيانات تعمل كدليل لمصداقيتها.

أما اذا لم تتوفر البيانات داخل المؤسسة أو خارجها، فعلى الباحث أن يعمل على توفير بيانات جديدة أو أولية إما عن طريق البريد أو التلفون أو المقابلة الشخصية أو الملاحظة أو التجارب أو المحاكاة. والخطوات اللاحقة في عملية البحث تتعلق بجمع البيانات من خلال هذه المصادر.

٥ ـ أدوات جمع البيانات.

يتوجب على الشخص الباحث وخلال اعداد أداة جمع البيانات ايجاد ربط فعّال ما بين المعلومات المطلوبة والأسئلة التي يجب ان تُسأل أو الملاحظات التي يجب تدوينها. ان نجاح أي دراسة يعتمد على مهارة الباحث ومقدرته الابداعية في ايجاد مثل هذا النوع من الربط. كما ان هذه المسؤولية تقع على عاتق الشخص الباحث وبالكامل.

٦ ـ تصميم العينة.

اذا لم يكن بالامكان دراسة جميع أفراد مجتمع الدراسة، فعلى الباحث القيام بعملية تصميم العينة والتي تهتم بمعرفة عناصر العينة. وحتى يتمكن الباحث من القيام بذلك فانه لا بد من تحديد مجتمع الدراسة والذي يجب أن نختار العينة منه بدقة. واختيار العينة يتطلب معرفة الطرق المختلفة والتي تستخدم في عملية الاختيار والتي يمكن تصنيفها الى طرق احتمالية وطرق غير احتمالية. ثم يتوجب على الشخص الباحث تحديد حجم العينة المناسبة.

٧ ـ جمع البيانات.

تعتبر عملية جمع البيانات مهمة جداً لأنها غالباً ما تتحكم بطبيعة نتائج البحث ومقدار الخطأ فيها كما انها تحتاج الى جزء كبير من ميزانية البحث. لذلك فان اختيار الأشخاص الباحثين وتدريبهم ومراقبتهم أمر ضروري لضمان عمل أبحاث فعّالة.

٨ ـ اعداد البيانات.

تبدأ عملية اعداد البيانات بعد تسجيلها. وهذا يتضمن أعمال التدقيق

والترقيم. التدقيق يتضمن مراجعة أداة جمع البيانات من حيث مؤهليتها في جمع البيانات والتوازن والكمال بينما الترقيم يتضمن تقسيم المستجوبين الى مجموعات بحيث يمكن تمثيلها بواسطة أرقام. وبعد ذلك تصبح البيانات جاهزة لوضعها في جداول من أجل اجراء التجاليل اللازمة اما بالطريقة اليدوية أو الآلية.

٩ ـ تحليل البيانات.

انه من المهم جداً لأن يكون التحليل للبيانات منسجم مع متطلبات المعلومات المطلوبة في الخطوة الثانية. وهذا يتطلب عادة استخدام النموذج أو الاختبار الاحصائي أو الرياضي والذي يتلاءم مع طبيعة أهداف الدراسة.

١٠ ـ عرض النتائج.

غالباً ما تعرض نتائج البحث على مدير المؤسسة أو الجهة المقدم لها البحث من خلال تقرير مكتوب أو عرض شفهي. كما يفضل أن تعرض نتائج البحث بطريقة مبسطة بحيث يتم معرفة المعلومات المطلوبة وفهمها بالنسبة لمتخذ القرار لاستخدامها في الحياة العملية. ففائدة البحث تكمن في طريقة عرض النتائج بغض النظر عن الفعالية التي تم انجاز الخطوات السابقة بها.

ـ الأخطاء في البحث Errors in Research

يمكن وقوع الأخطاء عند كل خطوة من خطوات البحث والذي بدوره قد يؤدي الى تزويد المدراء بمعلومات مضللة. وعليه فان مراقبة هذه الأخطاء غاية في الأهمية بالنسبة لمتخذي القرارات أو المدراء أو المؤسسات التي تدعم البحث. ويوجد هناك نوعان من الأخطاء (١) أخطاء العينة Sampling Errors (٢) أخطاء غير العينة Non-Sampling Errors

١ ـ أخطاء العينة Sampling Errors

تستخدم معظم الأبحاث التي تجري في هذه الأيام عينات من الأشخاص أو السلع أو المحلات أو الشركات. وبناءاً على نتائج العينة يقوم الباحث أو متخذ القرار بعمل استنتاجات عن مجتمع الدراسة والذي منه تم اختيار العينة. مثال : ان اتجاه المواطن الأردني نحو الصناعة المحلية قد أُشتقت من عينة من منطقة عمان. بما أن الباحث استخدم العينة لتقدير مجتمع الدراسة، فان هذا يؤدي الى وجود فروق ما بين نتائج العينة والنتائج الحقيقية لمجتمع الدراسة. وغالباً ما يطلق على هذا الفرق في النتائج بخطأ العينة.

٢ ـ أخطاء غير العينة Nonsampling Errors

تمثل اخطاء غير العينة جميع الاخطاء التي يمكن إن تقع في عملية البحث ما عدا اخطاء العينة. وهذا يتضمن ببساطة جميع النواحي بعملية البحث والتي تؤدي الى وقوع اخطاء مقصودة أو غير مقصودة. وبما ان هذه الأخطاء في عملية البحث تقع وبشكل مستمر فلا بد من ادراك ما يلي :

(١) ماذا يمكن للأخطاء غير العينية أن تحدث.

(٢) ما هو تاثير هذه الأخطاء على نتائج عملية البحث.

(٣) ما هي الخطوات الواجب اتباعها للتقليل من هذه الأخطاء.

ـ تأثير الأخطاء غير العينية The Effect of Nonsampling Errors

تتميز خصائص الأخطاء العينية بما يلي : (١) قابلة للقياس. (٢) تتناقص كلما زاد حجم العينة. لسوء الحظ، ان الأخطاء غير العينية صعبة القياس لا تتناقص مع ازدياد حجم العينة، بل تتزايد مع تزايد حجم العينة. ان تاثير الأخطاء غير العينية يجعل نتائج عملية البحث مضللة في اتجاهات ومقادير غير معروفة.

ـ أنوع الأخطاء غير العينية Types of Nonsampling Errors

١ ـ تعريف المشكلة الخاطيء Faulty Problem Definition

اذا لم يتمكن الباحث من تحديد المشكلة وبشكل دقيق، فان اي بحث يقوم به الشخص الباحث وبغض النظر عن تصميمه لا يخدم المشكلة الحقيقية. وعليه فان هذا سوف يتبعه خطأ في نتائج البحث وعدم امكانية في تطبيقها في حل المشكلة الحقيقية.

٢ ـ تحديد مجتمع الدراسة الخاطيء Defective Population Definition

يجب تحديد مجتمع الدراسة وبدقة ليتناسب مع أهداف الدراسة. فاذا كان مجتمع الدراسة هو طلبة جامعة مؤتة في كلية الاقتصاد والعلوم الادارية. وقد تم اختيار عينة من طلبة جامعة مؤتة كلية العلوم، فان هذا سوف يؤدي الى تحيّز في نتائج الدراسة. وعليه يكون نتائج الدراسة قابلة للتساؤل.

٣ ـ اطار الدراسة غير ممثل للمجتمع Frame is Nonrepresentative of the Population

يجب أن يكون اطار اختيار العينة مطابق لمجتمع الدراسة. مثال : تريد شركة استثمارات اختيار عينة أصحاب الأسهم باستخدام دليل الهاتف. ان اطار اختيار العينة قد لا تمثل مجتمع الدراسة لأن بعض حملة الأسهم قد لا يمتلكون خدمة الهاتف، لهذا فان نتائج الدراسة قابلة للشك.

٤ ـ أخطاء عدم الاستجابة Nonresponse Errors

تنتج بعض الأخطاء نتيجة رفض بعض أعضاء العينة المختارون ليكونوا أحد أعضاء العينة أو انشغالهم اثناء فترة مقابلتهم. فالعينة عند اختيارها كانت ممثلة، ولكن

عند عدم مقدرة مقابلة اي فرد من أفراد العينة يجعلها عينة غير ممثلة. وعليه فانه يطلق على هذا النوع من الأخطاء باخطاء عدم الاستجابة.

٥ ـ أخطاء في القياس Measurement Errors

القياس هي عملية تعيين أرقام للظواهر الملاحظة. فاذا قام الباحث بهذه العملية بطريقة غير دقيقة، فان هذا يؤدي الى حدوث أخطاء في القياس.

٦ ـ التصميم الرديء للاستبيان Poor Questinnaire Desigh

اذا كان تصميم الاستبيان غير فعّال اما بسبب عدم وضوح في كلمات جمل الاستبيان أو ازدواجية المعاني أوْ عدم تسلسل الأسئلة المنطقي، فان هذا سوف يؤدي الى تحيز في تعبئة الاستبيان من قبل الشخص المستجوب. وعليه فان البيانات المجمعّة عن طريق الاستبيان يكون مشكوك فيها.

٧ ـ أخطاء في اعداد البيانات Data Processing Errors

ان اعداد البيانات لمرحلة التحليل يتطلب تفريغ الاستبيان للتحضير لهذه المرحلة. وعملية الاعداد تتضمن اعطاء الاستجابات أرقام معينة ومن ثم تسجيلها ومن ثم تفريغها على أوراق خاصة بجهاز الحاسوب ... الخ. وانه من الممكن الوقوع باخطاء اثناء القيام بهذه العمليات أو تحيّز في عملية اعطاء الأرقام. وهذا يؤدي بالنهاية الى تحيز في نتائج دراسة البحث.

٨ ـ أخطاء في عملية التحليل Data Analysis Errors

انه لمن السهل جداً الوقوع في أخطاء اثناء عملية التحليل مثل جمع الأرقام الخاطيء. ولكن معظم الأخطاء اكثر تعقيد من ذلك. فمعظمها يتضمن تطبيق خاطيء لاجراءات تحليل البيانات.

٩ ـ أخطاء في تفسير النتائج Interpretation Errors

ان بعض الباحثين أو الأفراد يقومون بتفسير مجموعة من البيانات بشكل يتلائم مع اغراضهم. وهذا التفسير المتحيز قد يكون مقصود أو غير مقصود أو قد يكون بسبب عدم فهم نتائج دراسة البحث. ان هذا النوع من التفسير يثير الشكوك حول نتائج البحث.

وعليه فان :

الخطأ الكلي = الخطأ الناتج عن العينة + الخطأ الناتج عن غير العينة.

أسئلة

س١ : عرف البحث العلمي ؟

س٢ : أذكر أنواع البحث العلمي ثم قارن بينهما ؟

س٣ : حدد وبالتسلسل خطوات عملية البحث.

س٤ : لماذا يعتبر ضرورياً توقع خطوات عملية البحث من قبل الشخص الباحث ؟

س٥ : ما هو المقصود باخطاء غير العينة ؟

س٦ : ما هي خصائص أخطاء غير العينة ؟

س٧ : عرف جميع أنواع أخطاء غير العينة مع اعطاء مثال على كل نوع .

الفصل الثاني

تصميم البحث ومصادر البيانات

Research Design & Data Sources

تصميم البحث ومصادر البيانات
Research Design & Data Sources

ـ تصميم البحث Research Design

يعتبر تصميم البحث الخطة الأساسية والتي تعمل كدليل في جمع البيانات وتحليلها في مشروع البحث. ويمثل تصميم البحث اطار العمل الذي يحدد نوع المعلومات المطلوبة ومصادر البيانات واجراءات تجميعها. والتصميم الجيد يضمن ان المعلومات المجمعة تتوافق وأهداف الدراسة. كما يضمن التصميم الجيد جمع البيانات بطريقة دقيقة وباجراءات اقتصادية. ولا يوجد هناك تصميم بحث معياري أو مثالي يمكن ان يعمل كدليل للشخص الباحث لان هناك أكثر من تصميم بحث يمكن ان يحقق نفس نتائج البحث.

ـ أنواع الأبحاث Types of Research

ان هدف مشروع البحث هو الذي يحدد الخصائص المرغوب توفرها في تصميم البحث. ان أهداف مشروع البحث تعتمد على مراحل عملية القرار والتي نحتاج المعلومات لها. ويوجد هناك ثلاثة أنواع من الأبحاث بهذا الخصوص وهي :

١) أبحاث استكشافية Exploratory

٢) أبحاث نهائية Conclusive Research

٣) أبحاث مراقبة الاداء Performance-Monitoring Research

١ ـ الأبحاث الاستكشافية

تعتبر الأبحاث الاستكشافية مناسبة اذا كانت أهداف البحث تتضمن ما يلي :

(١) تحديد المشاكل أو الفرص.

(٢) تطوير صياغة دقيقة جديدة لفرصة أو مشكلة محددة بشكل واضح.

(٣) تحقيق منظور مستقبلي مناسب لمجموعة متغيرات تؤثر في مشكلة ما.

(٤) تحديد الأولويات بالنسبة للمشاكل أو الفرص.

(٥) ايجاد منظور مستقبلي اداري وبحثي بالنسبة لخصائص مشكلة البحث.

(٦) تحديد وتشكيل البدائل وسائل العمل المختلفة.

(٧) تجميع المعلومات عن المشاكل التي يمكن ان ترافق تنفيذ البحث النهائي.

تعتبر الأبحاث الاستكشافية الخطوة الأساسية لسلسلة من الدراسات المصمّمة لتزويد معلومات تساعد في اتخاذ القرارات. فالغرض من هذا النوع من البحث هو تشكيل الفرضيات بالنسبة لمشاكل فعلية أو فرص حقيقية يمكن إن تظهر عند وضع القرار. والمقصود بالفرضية جملة قابلة للرفض أو القبول تبحث في العلاقة بين متغيرين أو أكثر. ويجب ان تتضمن الفرضية مؤشر واضح في قياس المتغيرات واختيار العلاقة الموضوعة. ويمكن ان تكون الأبحاث الاستكشافية في الحالات التي تكون فيها الادارة تستجيب لأغراض تناقص الحصة السوقية المحلية للمؤسسة من خلال طرح السؤال التالي، ما هي المشكلة؟ فمهمة البحث الاستكشافي في هذه الحالة هي صياغة فرضية أساسية بحيث تهتم بسبب التناقص. ويمكن اعادة صياغة الفرضية الأساسية بفرضيات أخرى أكثر تحديد حتى نستطيع تحديد المشاكل أو الفرص بجمله. وتعتبر هذه الجملة الفرضية (الفرضيات) الأساسية بالنسبة للأسباب المؤثرة في حالة القرار. ويمكن اختبار هذه الفرضية أو الفرضيات في مرحلة لاحقة من عملية القرار بطرق البحث النهائية.

ومن الأمثلة على الفرضيات التي يمكن تطويرها في البحث الاستكشافي ما يلي :

١ ـ التركيز على القيمة الغذائية للسلعة في الحملة الدعائية يزيد من ادراك السلعة أكثر من التركيز على المذاق الجيد للسلعة.

٢ ـ تغيير محتوى المنتج من الشوكولاتة الصناعية مثلاً الى الشوكولاتة الحقيقية يزيد من تفضيل السلعة مقارنة مع السلع المنافسة.

٣ ـ تخفيض سعر التجزئة للمنتج بنسبة ١٠٪ سوف يؤدي الى زيادة في الحصة السوقية المحلية بنسبة ٢٪ خلال سنة.

إن البحث الاستكشافي يتعامل مع حالة قرار تكون فيه المعلومات المتوفرة قليلة. لذلك يجب أن يتميز تصميم البحث بالمرونة حتى يتمكن من استيعاب أي شيء غير متوقع لاكتشاف افكار لم تكن مدركة من قبل. ومن الاجراءات التي يمكن أن تساعد في عمل الأبحاث الاستكشافية ما يلي :

(١) البحث في المصادر الثانوية.

(٢) مقابلة الأشخاص ذوي المعرفة.

(٣) دراسة الحالات الماضية.

٢ ـ البحث النهائي

ان الهدف من تصميم البحث النهائي هو تزويد المعلومات لتقييم طرق العمل المختلفة للمساعدة في اختيار الطريقة الأمثل في حل مشكلة البحث. ويمكن تقسيم البحث النهائي الى قسمين :

أ ـ البحث الوصفي. **Descriptive Research**

ب ـ البحث السببي. **Causal Research**

أ ـ البحث الوصفي

تتضمن معظم دراسات الأبحاث الادارية ابحاثاً وصفية. تعتمد هذه الأبحاث وبشكل كبير على ما يقوله الشخص المستجوب والبيانات المتوفرة من مصادر البيانات الأولية. وتعتبر الأبحاث الوصفية ملائمة اذا كانت أهداف الدراسة تتضمن ما يلي :

١ ـ توضيح خصائص ظاهرة ادارية وتحديد درجة حدوثها.

٢ ـ تحديد درجة تأثير المتغيرات الادارية على الظاهرة.

٣ ـ عمل التنبؤات بالنسبة لحدوث الظاهرة الادارية.

يمكن للبحث الوصفي ان يحدد خصائص الظاهرة الادارية ويبيّن العلاقات ما بين المتغيرات، ولكن لا يمكنه من صياغة جمله بالنسبة للعلاقات السببية والتأثيرية. كما يمكن للشخص متخذ القرار ان يقوم بالتنبؤ بان بعض الأعمال المحددة سوف تؤدي بعض النتائج المحددة باستخدام البحث الوصفي ولكن هذا الدليل بعينه لا يكفي لتوضيح العلاقة السببية والتأثيرية.

ان خصائص البحث الوصفي والغرض منها تختلف اختلافاً فعلياً عن البحث الاستكشافي. يتميز البحث الوصفي الفعّال بوجود مشكلة واضحة واهداف محددة وتكون المعلومات المطلوبة مفصلة. كما يتصف تصميم البحث الوصفي بالتخطيط الجيد والهيكل المناسب لضمان نتائج دقيقة بنسبة عالية لان الغرض منه تزويد معلومات بالنسبة لأسئلة أو فرضيات محددة. والمقصود بالنتائج الدقيقة هو التقليل من الأخطاء التنظيمية وزيادة المصداقية بالشواهد او الأدلة المجمعة. وتشير الأخطاء التنظيمية الى التحيز الناتج من عملية القياس بينما المصداقية الى الحد الذي تكون فيه عملية القياس خالية من الأخطاء العشوائية وغالباً ما يستخدم البحث الوصفي أسلوب اختيار العينة من مجتمع الدراسة في وقت معين. ويعتبر هذا النوع من تصميم البحث اكثر الأنواع استخداماً واكثرها

معرفة عند الأشخاص. ويعتبر أسلوب تصميم العينة مفيداً في وصف خصائص مجتمع الدراسة وكذلك في تحديد درجة تكرار الظاهرة الادارية. مع العلم بأن أسلوب تصميم العينة غالباً ما يكون مكلفاً ويحتاج الى مهارات عالية من الشخص الباحث حتى يتمكن من تنفيذه بكفاءة.

ب ـ البحث السببي

ان عملية اتخاذ القرارات تتطلب افتراضات بالنسبة للعلاقات السببية والتأثيرية التي تظهر في النظام الاداري. ولتأكيد أو رفض هذه العلاقات فانه لا بد من تصميم بحث سببي لتجميع الأدلة بالنسبة لهذه العلاقات. ان هذا يتطلب القيام بتصميم بحث سببي يتصف بالتخطيط الجيد والهيكل المناسب للتقليل من تأثير الأخطاء التنظيمية وزيادة المصداقية وبالاضافة الى السماح بتزويد نتائج واضحة بالنسبة للعلاقات السببية.

وتعتبر الأبحاث السببية مناسبة في الأبحاث التي تتضمن الأهداف التالية :

١ ـ لفهم أي المتغيرات في الدراسة والتي سبّبت في حدوث ما تمّ التنبؤ به. فالتركيز هنا يكون في معرفة الأسباب التي أدت الى حدوث الأشياء.

٢ ـ لفهم طبيعة العلاقات الوظيفية ما بين العوامل المسبّبة وما هو التأثير المتوقع.

ان مصادر البيانات الرئيسية للبحث السببي هي :

(١) استجواب ذوي العلاقة. (٢) تنفيذ التجارب.

اذا كان بالامكان تحديد درجة العلاقة ما بين المتغيرات واختبار الفرضيات باستخدام طريقة الاستجواب، فانه ليس بالامكان التمييز بين المسببات كما هو الحال في طريقة التجارب. لذلك فان التصميم الجيد للتجربة يمكن ان يضمن أدلة ملاءمة وواضحة في التفسير بالنسبة للمسببات.

٣ ـ أبحاث مراقبة الأداء.

تزود أبحاث مراقبة الأداء معلومات بالنسبة لمراقبة النظام الاداري. وتعتبر هذه الأبحاث عنصر ضروري في مراقبة البرامج الادارية بالنسبة للتخطيط. والغرض من عمل مثل هذا النوع من البحث هو اعلام الادارة اما بوجود مشكل حقيقية أو فرص ضائعة.

ان أهداف أبحاث مراقبة الأداء هو لمراقبة وكتابة التغيّر :

(١) في قياس الانجاز مثل المبيعات والانتاج والحصة السوقية لتحديد ما اذا كانت الخطط تحقق الأهداف المرغوب فيها.

(٢) في الأهداف الفرعية مثل الادارك ومستوى المعرفة وعدد وحدات الانتاج ومستوى الأسعار بالنسبة للمواد الأولية لتحديد ما اذا تم تنفيذ البرامج الادارية حسب الخطط الموضوعة.

(٣) في المتغيرات البيئية مثل نشاطات المنافسة والتقدم التقني والظروف الاقتصادية لتحديد ماذا تمّ توقّع الظروف البيئية عند تشكيل الخطط.

ان مصادر البيانات المناسبة لأبحاث مراقبة الأداء تتضمن ما يلي :

(١) استجواب الأشخاص ذوي العلاقة.

(٢) البيانات الثانوية.

(٣) الملاحظة.

ـ مصادر البيانات. Data Source

يوجد هناك أربعة مصادر رئيسية للأبحاث الادارية، وهذه المصادر هي :

(١) المستجوبين ذوي العلاقة

(٢) المناظرة

(٣) التجارب

(٤) البيانات الثانوية.

١ ـ المستجوبين Respondents

يمثل المستجوبين مصدر رئيسي في تجميع البيانات الادارية. ويوجد هناك طريقتان رئيسيتان لتجميع االبيانات من المستجوبين وهي الاتصال والملاحظة. تتطلب طريقة الاتصال التفاعل بطريقة مباشرة من المستجوب في جمع البيانات. بينما لا تتطلب طريقة الملاحظة التفاعل المباشر من الشخص المراقب، بل هي عملية ادراك وتسجيل البيانات الملائمة عن حدث ما. ويمكن تجميع بيانات عن المستجوبين بطريقتين مختلفتين هي :

أ ـ طريقة الاتصال. Communication with Respondents

ب ـ طريقة الملاحظة. Observation of Respondents

أ ـ طريقة الاتصال :

تعتبر هذه الطريقة من أكثر طرق استخداماً في تزويد البيانات الادارية. وتعتبر هذه الطريقة منطقية في تجميع البيانات عن طريق طرح الأسئلة على مجموعة من الأشخاص. وتستخدم هذه الطريقة بشكل مستمر ويومي من قبل جميع الناس عن طريق طرح الأسئلة لأي شخص يعتقدون انه ذوي معرفة. وتعتبر هذه الطريقة مهمة في تجميع البيانات عندما تحتاج دراسة البحث بيانات عن الاتجاهات او الادراك او الحوافز أو المعرفة او سلوك مقصود. ويمكن ان يكون المستجوب مستهلك، عامل، مدير، فني، تاجر تجزئة او اي شخص ذو خبرة يمكن أن يعطي بيانات مفيدة لحالة القرار. وتتطلب عملية الاتصال الفعالة مع المستجوبين تدريب خاص ومهارات عالية اذا كانت البيانات مهمة جداً لحالة القرار، لأن البيانات تكون مضللة اذا كانت الأسئلة متحيزة أو اذا كان الأشخاص المستجوبين لا يمتلكون البيانات المرغوب فيها.

ان الأدوات المستخدمة في جمع البيانات عن طريق الأسئلة من المستجوبين هي :

(١) المقابلة الشخصية Personal Interview

(٢) المقابلة الهاتفية Telephone Interview

(٣) الاستبيان Questionnaire

ويمكن طرح الأسئلة على مجموعة قليلة من الأشخاص ذوي المعرفة كما هو الحال في الأبحاث النوعية (Qualitative Research) أو عمل دراسة مسحية تتضمن مئات الأشخاص كما هو الحال في الأبحاث الكمية (Quantitative Reasearch) .

ب ـ طريقة الملاحظة :

الملاحظة هي عملية معرفة وتسجيل بيانات ملائمة على حدث معين. تستخدم هذه الطريقة بشكل كبير في حياتنا اليومية. وتساعد طريقة الملاحظة في تجميع بيانات ذات علاقة مع مشكلة دراسة البحث في لحظة وقوعها. لذلك تمنع هذه الطريقة أو تقلل من الأخطاء الناتجة عن مقاطعة المستجوبين. ولكن لا نستطيع استخدام هذه الطريقة لتحديد الحوافز أو الاتجاهات أو الاعتقاد أو الشعور أو السبب من وراء سلوك معين.

٢ ـ المناظرة

ان الطريقة المنطقية لدراسة حالة القرار هو من خلال فحص وتحليل حالات مماثلة. وتتضمن الحالات المماثلة دراسة ما يلي :

(أ) الحالات التاريخية Case Histories

(ب) المحاكاة Simulation

أ ـ الحالات التاريخية

ان هذه الطريقة قديمة ومعروفة في دراسات العلوم الاجتماعية. ويتضمن تصميم البحث تحريّات مكثفة لتحديد الحالات ذات العلاقة مع مشكلة ودراسة البحث. ويقوم

مبدأ هذه الطريقة على اختيار مجموعة من الحالات التاريخية المماثلة وتحليلها للحصول على :

(١) تحديد المتغيرات المناسبة لمشكلة البحث .

(٢) معرفة طبيعة العلاقة بين المتغيرات .

(٣) تحديد طبيعة المشاكل أو الفرص التي تمّ دراستها في الحالات التاريخية .

وتعتبر هذه الطريقة مناسبة في الحالات التي تكون فيها المشكلة أو الفرصة ناتجة عن تفاعل معقد لسلسلة من المتغيرات . والحالات التي يمكن دراستها باستخدام دراسة الحالة التاريخية هي الحالات التي تهدف الى :

(١) مقارنة مستويات الانجاز مثل نوعية جيدة ورديئة .

(٢) تسلسل وقوع الأحداث .

ويمكن تجميع البيانات عن السجلات والتقارير أو ملاحظة بعض المتغيرات الرئيسية . ويتميز أسلوب البحث بالمرونة للاستفادة من وقوع الأشياء غير المتوقعة ولتطوير فهم أعمق لمشكلة البحث .

ب ـ المحاكاة

ان المقصود بالمحاكاة هو تطوير نموذج مماثل أو مشابه للظاهرة الحقيقية لاجراء الدراسة عليه دون التطرّق للظاهرة الطبيعية نفسها . ومن الأمثلة المعروفة عن المحاكاة هي نموذج الطائرات وخرائط الطرق .

أما بالنسبة للمحاكاة الادارية فهي عبارة عن تمثيل للنظام الاداري أو بعض مظاهره . وتعتبر هذه الطريقة حديثة نسبياً في تجميع البيانات وتعتمد على الحاسوب بشكل كبير . ويمكن استخدام أسلوب المحاكاة في الحصول على فهم أعمق لديناميكية النظام الاداري عن طريقة معالجة المتغيرات المستقلة وملاحظة تاثيرها على المتغيرات المتابعة . وتتطلب المحاكاة الادارية بيانات كمدخلات للنموذج تتعلق بخصائص الظاهرة

المُراد تمثيلها والعلاقات التي تربط ما بين هذه الخصائص.

من فوائد هذه الطريقة هو :

(١) قلّة التكاليف مقارنة مع طريقة مسح مجتمع الدراسة.

(٢) الوقت المطلوب لتجميع البيانات وتحليلها اقل.

(٣) السرّية في تنفيذ الدراسة. ومن سلبيات هذه الطريقة هو صعوبة تطوير نموذج محاكاة فعّال.

٣ ـ التجارب

تعتبر هذه الطريقة حديثة نسبياً في تجميع البيانات الادارية. وتنظيم البيانات المُجمعة من هذه الطريقة بطريقة معينة بحيث نستطيع الحصول على جملة واضحة بالنسبة لعلاقات السبب والتاثير.

وتستخدم هذه الطريقة في معالجة أو مراقبة المتغيرات المستقلة بدقة ويكون التاثير على المتغير أو المتغيرات التابعة قابل للقياس. فالهدف من طريقة التجربة هو مقياس تأثير المتغيرات المستقلة على المتغير التابع مع مراقبة المتغيرات الأخرى التي يمكن تقلل من مقدرة الباحث من عمل استنتاجات سببية فعالة.

٤ ـ البيانات الثانوية

يوجد هناك نوعان من البيانات هما الأولية والثانوية. البيانات الأولية هي البيانات التي تجمع لغايات البحث قيد الدراسة. اما البيانات الثانوية فهي البيانات المنشورة والموجودة اصلاً لغير غايات البحث قيد الدراسة. ويمكن الحصول على البيانات الثانوية من مصادر داخلية أو خارجية. وتعتبر البيانات التي تتوفر داخل المؤسسة التي تدعم البحث مصدراً داخلياً. أما البيانات التي نحصل عليها من خارج المؤسسة التي تدعم البحث مصادر خارجية.

إن الفائدة الرئيسية للبيانات الثانوية هو التوفير في الوقت والمال مقارنة مع مصادر البيانات الأولية. لذلك فانه من المهم جداً البحث في مصادر البيانات الثانوية أولاً قبل البدء بالحصول على البيانات الأولية. على الرغم من قلّة الحالات التي تكون فيها البينات الثانوية كافية لتحقيق متطلبات البحث قيد الدراسة، إلا إنها تساعد في :

(١) صياغة مشكلة البحث.

(٢) تساعد في تحديد طرق جمع البيانات وأنواعها والتي تساعد في تحقيق المعلومات المطلوبة.

(٣) تعمل كمصدر لمقارنة البيانات والتي من خلالها نستطيع تفسير وتقييم البيانات الأولية.

(٤) يمكن أن تساعد في ايجاد الطريقة الأمثل لمعالجة البيانات الأولية ذات العلاقة بمشكلة البحث قيد الدراسة.

(٥) تساهم في تعميق فهم مشكلة البحث عند الشخص الباحث.

ومن المساوىء الرئيسية للبيانات الثانوية تتعلق فيما يلي :

(١) مدى ملاءمة البيانات الثانوية للمعلومات المطلوبة عن مشكلة البحث.

(٢) دقة البيانات الثانوية المتوفرة .

وعلى الرغم من أهمية البيانات الثانوية للشخص الباحث، الا انها لا تكفي في معظم الحالات لحل مشكلة البحث قيد الدراسة لعدم وجود المعلومات الدقيقة والكافية عن مشكلة البحث خاصة اذا لم تناقش مشكلة البحث من قبل. لذلك يتوجب على الشخص الباحث الحصول على البيانات الأولية على الرغم من ارتفاع تكاليف الحصول عليها والوقت الطويل اللازم لتجميع البيانات الأولية وتفريغها وتبويبها ومن ثم تحليلها.

أسئلة

س١ : ناقش طبيعة ودور تصميم البحث في الأبحاث الادارية.

س٢ : ما هو نوع تصميم البحث المتعلق بالأبحاث الاستكشافية ؟

س٣ : لماذا غالباً ما يستخدم البحث الاستكشافي في المراحل الأولية من عملية البحث؟

س٤ : ما هو المقصود بالبحث الوصفي ؟

س٥ : كيف يختلف تصميم البحث في الأبحاث الوصفية مقارنة مع الأبحاث الاستكشافية ؟

س٦ : ما هي أهداف أبحاث مراقبة الأداء ؟

س٧ : ناقش المصادر الرئيسية في جمع البيانات الأولية ؟

س٨ : أذكر الفرق الرئيسي ما بين الأبحاث الكمية والأبحاث النوعية ؟

س٩ أذكر فوائد وسلبيات المحاكاة ؟

الفصل الثالث

المعاينة (العينات)

Sampling

المعاينة (العينات)
Sampling

- مقدمة :

تستخدم المعاينة (العينة) بشكل مستمر في الأبحاث الادارية لمقدرتها في توفير الفوائد الرئيسية التالية مقارنة مع دراسة جميع افراد مجتمع الدراسة.

١ - توفير بالمال.

اذا كان مجتمع الدراسة كبير، فان تكاليف تجميع البيانات بواسطة المقابلة الشخصية سوف تكون عالية اذا ما تمّ مقابلة جميع افراد الدراسة. وعليه فان اختيار عينة من مجتمع الدراسة سوف يوفر بالتكاليف اللازمة لتجميع البيانات.

٢ - توفر بالوقت.

ان اختيار عينة من مجتمع الدراسة سوف يؤدي الى تقليل الوقت اللازم لتجميع البيانات من افراد العينة مقارنة مع جميع افراد مجتمع الدراسة.

٣ - ممكن ان تكون بيانات العينة اكثر دقة احياناً.

قد تبدو هذه الفائدة مدهشة في البداية، ولكن هذا بالواقع صحيح. ،ينتج ذلك من مصادر متعددة والتي يطلق عليها اخطاء غير العينة والتي تظهر في عملية البحث. ان دراسة مجتمع الدراسة تتطلب مقابلات اكثر وعدد اكبر من المشرفين على المقابلات وعدد كبير لتفريغ البيانات واعدادها للتحليل ... الخ. فكلما كانت الدراسة صغيرة كلما زاد احتمال الحصول على اشخاص متخصصين وذوي مهارات لكل مرحلة من مراحل البحث. وكلما زاد عدد الأشخاص كلما قلت

نوعيتهم وخبرتهم وبالتالي تصبح المراقبة والاشراف على النشاطات المختلفة صعبة. كما ان دراسة مجتمع الدراسة تأخذ الوقت الكثير وعليه يمكن حدوث بعض التغيّر على الظاهرة الادارية قيد البحث. مثال : ان دراسة ادراك مجتمع الدراسة لسلعة جديدة مثلاً له معنى وفائدة في لحظة معينة لأن ادراك مجتمع الدراسة لهذه السلعة يزداد مع مرور الوقت. وعليه فان جميع ما ذ؛كر يؤدي الى الوقوع بأخطاء كثيرة ونتائج أقل دقة أو متحيزة.

وكما هو الحال بالنسبة لمجتمع الدراسة، فان دراسة العينة تتضمن أخطاء غير العينة ولكن بدرجة أقل. كما ان دراسة العينة تتضمن أخطاء العينة والتي تخلو منها دراسة المجتمع. تستخدم دراسة العينة لتقدير بعض القيم عن مجتمع الدراسة. فاذا كان هناك فرق ما بين القيم الناتجة عن دراسة العينة والقيم الحقيقية لمجتمع الدراسة، فان ذلك يؤدي الى تحقّق خطأ العينة.

وعليه تكون دراسة العينة أكثر دقة من دراسة مجتمع الدراسة اذا كان مجموع أخطاء العينة وغير العينة للعينة أقل من أخطاء غير العينة لمجتمع الدراسة.

٤ ـ تكون دراسة العينة أفضل من مجتمع الدراسة في الحالات التي تؤدي فيها الدراسة الى إتلاف أو تلويث للعنصر الذي وقعت عليه الدراسة. مثال : فحص مذاق سلعة معينة. ان فحص مذاق جميع الوحدات المنتجة من شوكولاتة عطا علي ليست الطريقة الصحيحة لعمل مشروع مربح.

ـ عملية اختيار العينة

بعد بيان أسباب دراسة العينة بدلاً من مجتمع الدراسة، فانه لا بد من طرح سؤال كيفية اختيار العينة. وقبل البدء في الاجابة على هذا السؤال فانه لا بد من تعريف بعض المصطلحات الخاصة بالمعاينة.

١ ـ العنصر Element

العنصر هو الوحدة التي تزود المعلومات. كما يعتبر العنصر الأساس في عملية التحليل الاحصائي الواجب عمله. ويمكن ان يكون العنصر الانسان أو السلعة أو العائلة. فشكل العنصر في أي عينة يعتمد على طبيعة مشكلة الدراسة وأهدافها.

٢ ـ المجتمع Population

المجتمع هو مجموعة جميع العناصر والتي تمّ تعريفها من قبل عملية اختيار العينة.

٣ ـ وحدة العينة Sampling Unit

وحدة العينة هي العنصر أو مجموعة العناصر المتوفرة في عملية اختيار العينة.

٤ ـ اطار العينة Sampling Frame

يمثل اطار العينة القائمة التي تحتوي على جميع وحدات العينة المتوفرة للاختيار عند عملية اختيار العينة.

٥ ـ مجتمع الدراسة Study Population

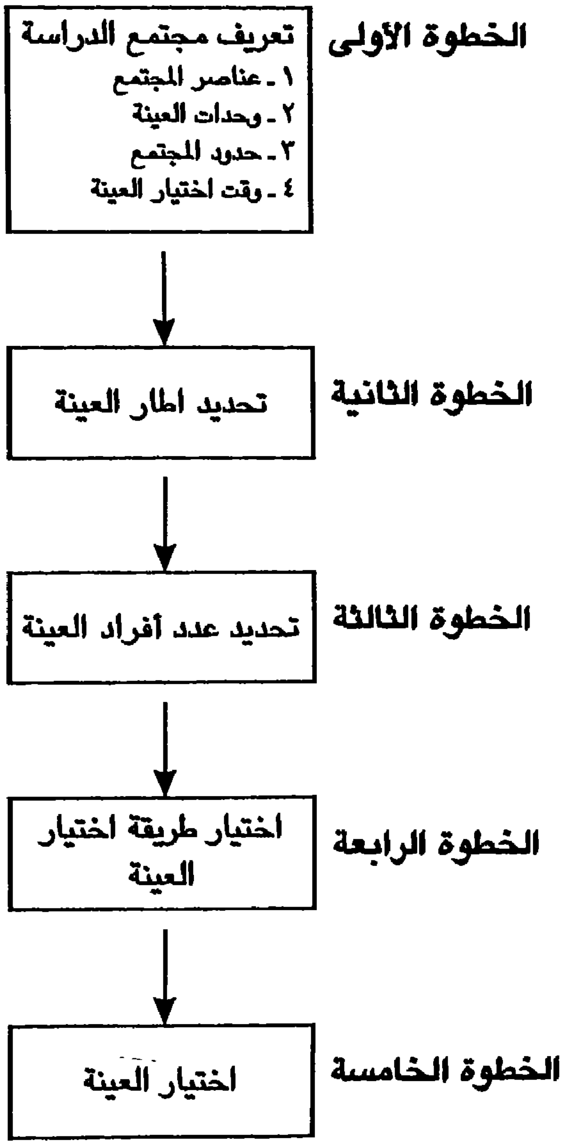
يمثل مجتمع الدراسة مجموعة العناصر والتي منها فعلاً تم اختيار العينة.

ـ خطوات اختيار العينة

يمثل الشكل (٣ ـ ١) خطوات اختيار العينة :

شكل (٣ - ١)

خطوات اختيار العينة



الخطوة الأولى : تعريف مجتمع الدراسة ويتضمن ذلك تحديد عناصر مجتمع الدراسة ووحدات العينة وحدود المجتمع ووقت اختيار العينة.

الخصوة الثانية : تحديد اطار العينة والذي من خلاله سوف يتم اختيار العينة.

الخطوة الثالثة : تحديد حجم العينة ويتضمن ذلك تحديد عدد العناصر التي يجب أن تشكل العينة. وسوف نقوم بذلك في فصل لاحق.

الخطوة الرابعة : اختيار اجراءات طريقة اختيار العينة وسوف يتم شرحه في الفصول القادمة.

الخطوة الخامسة : الاختيار الفعلي للعينة بناء على الاجراءات التي تم اختيارها في الخطوة الرابعة.

- اجراءات اختيار العينة Sampling Procedures

يوجد هناك العديد من الاجراءات والتي يستطيع الباحث من خلالها اختيار العينة المناسبة. ويمكن تقسيم العينات بشكل عام الى قسمين :

١ - عينة احتمالية **Probability Sample**

٢ - عينة غير احتمالية **Nonprobability Sample**

في العينات الاحتمالية، يكون كل عنصر من عناصر المجتمع له فرصة معلومة لأن يكون أحد أعضاء العينة. فتتم المعاينة الاحتمالية بالقوانين الرياضية التي لا تترك مجال لتدخل الشخص الباحث. لاحظ اننا قلنا فرصة معلومة وليس فرصة متساوية لأن يكون العنصر أحد أفراد العينة. ولا تتمتع عناصر مجتمع الدراسة بنفس الفرصة الاحتمالية لأن تكون أحد أفراد المجتمع إلا في حالة خاصة من المعاينة والتي تدعى بالعينة العشوائية البسيطة.

ان العينات الاحتمالية تمكّن الى حد ما من قياس الفرق ما بين قيم العينة عن قيم

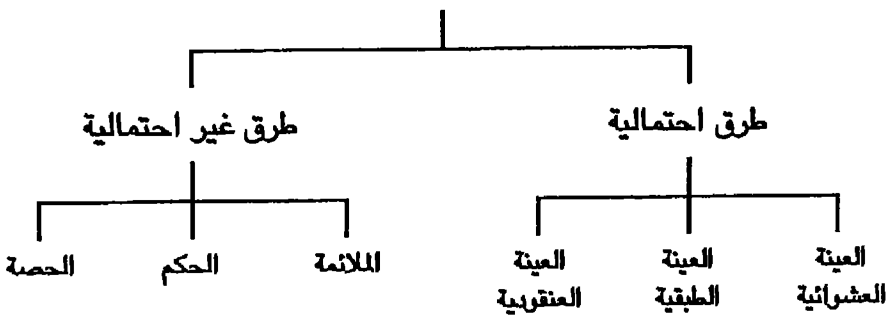
المجتمع قيد الدراسة. ويدعى هنا الفرق بالخطأ العيني.

أما في العينات اللاحتمالية، فان اختيار اي عنصر من عناصر مجتمع الدراسة كأحد أفراد العينة يعتمد في مرحلة من المراحل على الحكم الشخصي للباحث او الشخص الذي ينفذ عملية المقابلة. لذلك لا يمكن معرفة احتمال اختيار اي عنصر من عناصر مجتمع الدراسة لأن يكون احد أفراد العينة في العينات اللاحتمالية. وكنتيجة فاننا لا نستطيع حساب الخطأ العيني الذي يمكن ان يتحقق في هذه الحالة. وهذا يعني انه من الصعب جداً معرفة ما اذا كانت التقديرات الاحصائية المحسوبة من العينة دقيقة أم لا.

ويمكن استخدام اكثر من طريق لاختيار العينة العشوائية وكذلك نفس الشيء بالنسبة لاختيار العينة غير الاحتمالية. والشكل (٣ - ٢) يبين هذه الطرق الممكنة.

الشكل (٣ - ٢)

طرق اختيار العينة



تتكون الطرق الاحتمالية في اختيار العينة كما هو مبين في الشكل (٣ - ٢) مما يلي :

١ - العشوائية البسيطة Simple Random Sample

٢ - الطبقية Stratified Sample

٣ - العنقودية Cluster Sample

بينما تتكون الطرق اللاحتمالية في اختيار العينة من نفس الشكل مما يلي :

١ - الملائمة Convenience Sample

٢ - الحكم Judgement

٣ - الحصة Quota

في هذا الفصل سوف نقوم بشرح الطرق الثلاثة المختلفة في اختيار العينة غير الاحتمالية. بينما نخصص الفصلين اللاحقين لدراسة الطرق الثلاثة المختلفة في اختيار العينة الاحتمالية.

- طرق اختيار العينة اللاحتمالية

١ - طريقة الملائمة.

يقوم مبدأ عمل هذه الطريقة في اختيار العينة كما هو واضح من عنوان هذه الطريقة على أساس الملائمة بالنسبة للشخص الباحث. ومن الأمثلة على ذلك ما يلي :

١ - السؤال عن أشخاص متطوعين لفحص سلعة أو مجموعة من السلع ومن ثم استخدام هؤلاء الأشخاص كعينة.

٢ - إيقاف مجموعة من الأشخاص في مجمع تجاري لمعرفة رأيهم في موضوع معين كعينة.

٣ ـ استخدام طلبة جامعيين لاجراء تجارب كعينة.

٤ ـ استخدام مجموعة من الأشخاص تمّت مقابلتهم تلفزيونياً كعينة.

نلاحظ في جميع الأمثلة السابقة انه تمّ اختيار عنصر العينة سواء كان الاختيار من قبل الشخص الباحث أو تطوعاً من العنصر نفسه بناءاً على سهولة الحصول عليه. وفي جميع الحالات لم يكن معروف مجتمع الدراسة الذي تم اختيار العينة منه. بالاضافة الى ان احتمال اختيار كل عنصر من عناصر المجتمع لأن يكون احد افراد العينة غير معروف. ففي المثال الأول : الأشخاص المتطوعين فقط هم الذين تمّ اختيارهم للعينة واي شخص غير متطوع تمّ استثناؤه. وكذلك النفس الشيء بالنسبة للمثال الثاني. فالأشخاص الذين تصادف وجودهم في المجمع التجاري لحظة وجود الباحث تمّ اختيارهم لتشكيل العينة. اما الذين لم يكونوا موجودين في المجمع التجاري تمّ استبعادهم.

من الأمثلة السابقة، نجد ان الفروق ما بين قيم مجتمع الدراسة وقيم العينة من حيث الحجم والاتجاه غير معروفة. لذلك فانه من الصعب جداً قياس الخطأ الناتج عن العينة. وهذا يعني صعوبة استنتاج جمل محدّدة ذات معنى عن النتائج التي نحصل عليها من العينة. لذلك يمكن تبرير استخدام عينة الملائمة في مرحلة البحث الاستكشافي والتي تستخدم في تحديد الفرضيات للدراسة قيد البحث.

٢ ـ طريقة الحكم.

يقوم مبدأ عمل هذه الطريقة في اختيار العينة على أساس المعرفة في موضوع البحث قيد الدراسة. فالأشخاص الذين يمكن ان تكون لديهم المعرفة في الاجابة على اسئلة البحث قيد الدراسة هم الذين يشكلون عينة الدراسة. اما الأشخاص اللذين يتوقع عدم معرفتهم في موضوع البحث قيد الدراسة فيتم استبعادهم عن العينة. لذلك نجد ان احتمال اختيار اي عنصر من عناصر مجتمع الدراسة لأن يكون احد اعضاء مجتمع الدراسة باستخدام هذه الطريقة غير معروف. كما نجد ان الفروق ما بين قيم

مجتمع الدراسة وقيم العينة غير معروف. وهذا يعني ان درجة الخطأ في العينة واتجاهه غير معروف، وبالتالي صعوبة استنتاج جمل محددة ذات معنى عن النتائج التي نحصل عليها من العينة. ولكن غالباً من تكون الجمل الاستنتاجية عن نتائج العينة باستخدام طريقة الحكم افضل من الجمل الاستنتاجية عن نتائج العينة باستخدام طريقة الملائمة لأن افراد العينة غالباً ما يكونوا على علم بموضوع البحث قيد الدراسة. لذلك نجد ان استخدام هذه الطريقة في الحياة العملية افضل مقارنة مع طريقة الملائمة.

٣ ـ طريقة الحصة.

تعتبر طريقة الحصة في اختيار العينة حالة خاصة من طريقة الحكم في اختيار العينة. لذلك نجد ان الشخص الباحث يتخذ خطوات واضحة للحصول على عينة مشابهة لمجتمع الدراسة بناءاً على بعض الخصائص. وكمثال على ذلك، وجود تعليمات عند الشخص الباحث لاختيار نصف الأشخاص المراد مقابلتهم ممن هم في سن الثلاثين أو أكثر والنصف الآخر ممن هم دون سن الثلاثين. فقد تمّ في هذه الحالة استخدام عامل العمر كاساس في اختيار العينة. وهذا يعني ان الباحث يعرف مسبقاً ان مجتمع الدراسة يتكون من قسمين متساويين من الأشخاص الذي هم في سن الثلاثين وأكثر والذين هم دون سن الثلاثين. نلاحظ في هذا المثال ان خاصية العمر هي التي استخدمت في عملية اختيار العينة.

ولاختيار عينة أكثر واقعية لكي تكون ممثلة لمجتمع الدراسة، فانه لا بد من استخدام خاصية معينة لاستخدامها كخصائص مراقبة عند اختيار عناصر العينة. لذلك حتى نتمكن من اختيار عينة الحصة فانه لا بد من:

(١) تحديد قائمة بالخصائص التي يتوجب توفرها في عناصر عينة الحصة.

(٢) معرفة توزيع هذه الخصائص في مجتمع الدراسة.

مثال : افرض اننا نريد اختيار عينة الحصة بناء على خاصيتين ذات اهتمام هما العمر واللون وكما يلي :

١ ـ العمر : مجموعتين ـ تحت سن الثلاثين وسن الثلاثين أو أكثر.

٢ ـ اللون : مجموعتين ـ أبيض وغير أبيض.

نلاحظ هنا انه يوجد أربعة خلايا أو حصص يمكن الاختيار من بينهما هما :

١ ـ تحت سن الثلاثين واللون أبيض.

٢ ـ سن الثلاثين أو أكثر واللون ابيض.

٣ ـ تحت سن الثلاثين واللون غير أبيض.

٤ ـ سن الثلاثين أو أكثر واللون غير أبيض.

ان الشيء الذي يجب معرفته الآن هو نسبة مجتمع الدراسة في كل خلية أو حصة. ان صعوبة تحديد نسبة مجتمع الدراسة في كل حصة تزداد مع تزايد الخصائص المحددة في اختيار عينة الحصة. لاحظ ماذا سيحدث لعدد خلايا العينة اذا زاد عدد الخصائص المحددة في اختيار العينة وكذلك عدد المجموعات لكل خاصية. افرض وجود أربعة خصائص محددة للاختيار وكما يلي :

١ ـ العمر : ٤ مجموعات ـ (١) تحت ١٨ (٢) ١٨ ـ ٣٠ (٣) ٣١ ـ ٤٥ (٤) أكثر من ٤٥.

٢ ـ اللون : ٣ مجموعات ـ (١) أبيض (٢) أسود (٣) شيء آخر.

٣ ـ التعليم : ٤ مستويات ـ (١) مدرسة ابتدائية (٢) ثانوية (٣) كلية مجتمع (٤) جامعة.

٤ ـ الدخل : ٥ مجموعات ـ (١) تحت ١٠٠٠ دينار (٢) ١٠٠٠ ـ ٢٥٠٠ دينار (٣) ٢٥٠١ ـ ٥٠٠٠ دينار (٤) ٥٠٠١ ـ ٧٥٠٠ دينار (٥) أكثر من ٧٥٠١ دينار.

ان عدد الخلايا أو الحصص التي يمكن أن نختار من بينها عينة الحصة هو ٤ X ٣ X ٤ X ٥ = ٢٤٠ خلية. وحتى نتمكن من اختيار الخلية الأكثر تمثيل لمجتمع الدراسة فانه لا بد من معرفة معلومات عن نسبة مجتمع الدراسة في كل خلية من هذه الخلايا. وعليه فان هذا قد يكون من الصعب جداً بل من المستحيل عمله أو معرفته.

المشاكل التي تواجه عينة الحصة.

١ ـ يجب أن تكون نسبة مجتمع الدراسة في كل خلية من الخلايا دقيقة وحديثة وليست نسب قديمة وغالباً ما يكون هذا صعب بل مستحيل.

٢ ـ يجب تحديد جميع الخصائص المحددة المناسبة. بمعنى آخر وجوب اختيار جميع الخصائص التي لها علاقة في قياس الشيء موضوع البحث. مثال : اذا أردنا معرفة اتجاه الناس نحو الرجال ذوي الشعر الطويل، فانه من الخطأ جداً عدم استخدام عامل العمر كخاصية مراقبة لأن العمر له علاقة بالاتجاه نحو الشعر الطويل للرجال. فاذا تمّ استبعاد أي خاصية ذات علاقة بموضوع الدراسة سواء كان ذلك عن قصد أو غير قصد، فان ذلك يؤدي الى الحصول على نتائج مضللة.

٣ ـ ان المشكلة الثالثة التي ترافق استخدام طريقة الحصة في اختيار العينة هو وجود عدد كبير من خصائص المراقبة والتي تستخدم في اختيار العينة. لأن عدد الخلايا التي يجب أن يتعامل معها الشخص الباحث وكما تبين قبل قليل سوف يكون كبير. وبالتالي فان اختيار الخلية أو الحصة التي يكون درجة تمثيلها لمجتمع الدراسة جيد تكون عملية صعبة جداً.

٤ ـ المشكلة الرابعة التي تعاني منها هذه الطريقة هي عملية اختيار الباحث الفعلية لأعضاء العينة. ان عملية ايجاد الأشخاص المتوقع توفر الخصائص المرغوب بها فيهم قد تؤدي بالشخص الباحث الى استثناء الأشخاص العدوانيين أو الذين يسكنون في بيئات فقيرة أو الذين يعيشون في بيوت قديمة وما الى ذلك. ان هذا السلوك من

الشـخص الباحـث قد يؤدي الى تحـيز غـير مـقصـود وبالتالي يـنعكس على نتـائج الدراسة .

لذلك، تعتبر طريقة الحصة فعالة في مراحل البحث الأولية وغالباً ما تزود هذه الطريقة نتائج جيدة اذا ما تم استخدامها بحذر . ولكن نتائج هذه الطريقة غالباً ما تكون أقل دقة مقارنة مع النتائج التي يحصل عليها من العينات العشوائية . وبشكل عام تستخدم هذه الطريقة وبشكل واسع في الحياة العملية .

أسئلة

س١ : لماذا يستخدم أسلوب العينة في الأبحاث الأولية ؟

س٢ : ميّز ما بين المصطلحات التالية : العنصر، مجتمع الدراسة، وحدة العينة، اطار العينة، ومجتمع الدراسة.

س٣ : قارن ما بين العينات العشوائية وغير العشوائية.

س٤ : ما هي طبيعة الأخطاء التي يمكن أن تتحقق نتيجة استخدام اجراءات العينات غير الاحتمالية ؟

س٥ : قارن ما بين الطرق الثلاثة المختلفة في اجراءات اختيار العينة غير العشوائية.

س٦ : لماذا تستخدم اجراءات اختيار العينة غير الاحتمالية في الحياة العملية ؟

الفصل الرابع

العينة العشوائية البسيطة وحجم العينة

Simple Random Sampling and Sample Size

العينة العشوائية البسيطة وحجم العينة
Simple Random Sampling and Sample Size

لقد تم عرض المبادىء الضرورية لفهم المعاينة في الأبحاث الادارية في الفصل السابق. كما قد تمّ توضيح الفروق الرئيسية والأساسية ما بين اجراءات اختيار العينة بالطرق الاحتمالية والطرق غير الاحتمالية. وسوف يتم التركيز في هذا الفصل على شرح مفصل لاجراءات اختيار العينة باستخدام العشوائية البسيطة. كما ان هذا الفصل سوف يركز على موضوع دقة نتائج العينة. هذا بالاضافة الى تناول موضوع تحديد حجم العينة.

- العينة العشوائية البسيطة

تتميز طريقة اختيار العينة بالعشوائية البسيطة بالفرص المتساوية التي يتمتع بها جميع أفراد مجتمع الدراسة عند عملية اختيار عناصر العينة. كما ان احتمال اختيار أي عنصر بهذه الطريقة يكون مستقلاً عن اختيار العناصر الأخرى. لذلك تعتبر العينة العشوائية البسيطة من افضل العينات على الاطلاق في المقدرة على تمثيل مجتمع الدراسة من حيث توفر الخصائص. ان اختيار عناصر العينة العشوائية البسيطة يقوم على أساس جداول الأرقام العشوائية.

- تعريف بعض الرموز

حتى نستطيع البدء في دراسة العينة العشوائية البسيطة فانه لا بد من التعرف على تعريف بعض المصطلحات.

١ ـ المعلمة Parameter

المعلمة عبارة عن قيم تستخدم للقياس بالنسبة لمجتمع الدراسة قيد البحث. وتكون هذه القيم حقيقية اذا ما تمّ اجراء مسح شامل لمجتمع الدراسة قيد البحث اذا كانت الاجراءات خالية من أخطاء غير العينة. فمتوسط عمر طلاب مادة أساليب البحث العلمي عبارة عن معلمة. وكذلك متوسط دخلل موظفي جامعة مؤتة عبارة عن معلمة.

٢ ـ قيم احصائية Statistic

القيم الاحصائية عبارة عن قيم للقياس بالنسبة للعينة التي تم اختيارها.

وغالباً ما تستخدم القيم الاحصائية في تقدير معالم مجتمع الدراسة. لذلك يكون اعتبار متوسط عمر طلبة مادة أساليب البحث العلمي أو متوسط دخل موظفي جامعة مؤتة كقيم احصائية اذا ما تمّ قياسهم على أساس العينة.

٣ ـ بعض الرموز التي هي بحاجة الى تعريف Some Symbols to Get Straight

يوجد هنام نوع من التعارف في استخدام بعض الرموز لوصف معالم مجتمع الدراسة وبعض الرموز لوصف عينة الدراسة. ولقد تم الاصطلاح على استخدام الرموز الاغريقية لوصف معالم مجتمع الدراسة والرموز الانجليزية في وصف قيم العينة. والجدول (٤ ـ ١) يبين ذلك.

جدول (٤ - ١)

الرموز المستخدمة في وصف المجتمع وعينة الدراسة

قراءة من اليمين الى اليسار

انواع المقاييس	العلاقة	رموز مجتمع الدراسة	رموز العينة الدراسة
١) المقاييس المتصلة	الوسط او متوسط المتغير المتصل	μ	$\bar{X}$
	التباين للمتغير المتصل	σ^2	S^2
٢) المقاييس المنفصلة او مقاييس ذو الحدين (مثال : هل انت ذكر؟ نعم او لا)	نسبة من قال «نعم»	π	P
	نسبة من قال «لا»	$(\pi - 1)$	(P - 1)
	التباين للنسب	σ^2	S^2

مثال : يمثل جدول : (٤ - ٢) مجتمع الدراسة الذي سوف يستخدم في توضيح أفكار العينة. ويتكون مجتمع الدراسة من مجموعة الطلاب المسجلين لمادة أساليب البحث العلمي عند دكتور معين. وقد جمع الدكتور من الطلبة ثلاثة أنواع من المعلومات هي :

١ - أول منزلة عشرية في رقم الطالب الجامعي (عمود ١) .

٢ - عمر كل طالب في المادة (عمود ٢) .

٣ ـ جمله حول ما اذا كان الطالب يرغب بتسجيل مادة الاحصاء قبل التخرج (على فرض انها مادة اختيارية). وقد تمّ ترميز الاجابات كما يلي :

أ ـ نعم ـ يرغب الطالب في تسجيل مادة الاحصاء.

ب ـ لا ـ لا يرغب الطالب بتسجيل مادة الاحصاء (عمود ٣).

اما بالنسبة للأعمدة الأخرى فسوف تستخدم في الحسابات اللاحقة. وسوف تستخدم الخانة الأولى من الرقم الجامعي لتحديد كل عنصر من عناصر المجتمع. اما المتغيرات في العمودين الثاني والثالث سيكونان المتغيران ذات الاهتمام لعملية القياس. ويجب ان نذكر هنا ان العمر هو متصل بينما اختيار مادة الاحصاء هو متغير متقطع. دعنا نحدد المتغيرات كما يلي :

العمر = س ١

اختيار مادة الاحصاء = س ٢

ـ حساب معالم مجتمع الدراسة بالنسبة للمتغيرات المتصلة.

اننا سوف نقوم بحساب المعالم المحسوبة لمجتمع الدراسة كطريقة لتوضيح طبيعة قيم العينة الاحصائية. والجدول (٤ ـ ٢) يمثل مسح شامل للعمر واختيار مادة الاحصاء كمتطلب اختياري لمادة أساليب البحث العلمي.

ان متغير العمر هو متغير متصل. دعنا الآن نجد مقاييس النزعة المركزية ومقاييس التشتت لمتغير العمر. والذي يتم كما يلي :

مقاييس النزعة المركزية : الوسط أو المعدل = μ

مقاييس التشتت : الانحراف المعياري أو التباين = σ و $\sigma2$ على التوالي.

ان الوسط لمجتمع الدراسة ببساطة يساوي مجموع القيم مقسوماً على عدد عناصر المجتمع. وعليه

$$\mu = \frac{\sum_{s=1}^{\text{ن}} X_s}{\text{ن}}$$

$$\mu = \frac{\text{١٨} + \ldots + \text{٢٧} + \text{٢٥}}{\text{٥٠}}$$

$$\mu = \text{٢٣,٧}$$

وهذا يعني ان متوسط العمر في الصف هو ٢٣٫٧ سنة. أما بالنسبة للتباين فيساوي مربع مجموع انحراف البيانات عن وسطها الحسابي مقسوماً على عدد افراد المجتمع.

جدول (٤ - ٢)

مسح شامل للعمر واختيار مادة الاحصاء كمتطلب اختياري للطلبة.

(١) أول خانة في رقم الطالب الجامعي	(٢) العمر ($س_{١}$)	(٣) اختيار مادة الاحصاء ١ = نعم ٠ = لا ($س_{٢}$)	(٤) $(س_{١} - \mu)$	(٥) $(س_{١} - \mu)^{٢}$
١	٢٥	١	١ر٣	١ر٦٩
٢	٢٧	٠	٣ر٣	١٠ر٨٩
٣	٢٩	١	٥ر٣	٢٨ر٠٩
٤	٣١	١	٧ر٣	٥٣ر٢٩
٥	٢٥	٠	١ر٣	١ر٦٩
٦	٢٩	٠	٥ر٣	٢٨ر٠٩
٧	٢٧	٠	٣ر٣	١٠ر٨٩
٨	٢٤	٠	٠ر٣	٠ر٠٩
٩	٢٧	١	٣ر٣	١٠ر٨٩
١٠	٢٨	١	٤ر٣	١٨ر٤٩
١١	٣٣	٠	٩ر٣	٨٦ر٤٩
١٢	٢٩	١	٥ر٣	٢٨ر٠٩
١٣	٢٦	٠	٢ر٣	٥ر٢٩
١٤	٢٨	٠	٤ر٣	١٨ر٤٩
١٥	٢٨	١	٤ر٣	١٨ر٤٩
١٦	٢٦	٠	٢ر٣	٥ر٢٩
١٧	٢٦	١	٢ر٣	٥ر٢٩
١٨	٣٦	١	١٢ر٣	١٥١ر٢٩
١٩	٢٩	٠	٤ر٣	١٨ر٤٩
٢٠	٢٦	٠	٢ر٣	٥ر٢٩

طلبة دراسات عليا

تابع / جدول (٤ - ٢)

(١) أول خانة في رقم الطالب الجامعي	(٢) العمر ($س_1$)	(٣) اختيار مادة الاحصاء ١ = نعم ٠ = لا ($س_2$)	(٤) $(س_1 - \mu)$	(٥) $(س_1 - \mu)^2$
٢١	٢١	٠	– ٢ر٧	٧ر٢٩
٢٢	١٩	٠	– ٤ر٧	٢٢ر٠٩
٢٣	٢٤	٠	٠ر٣	٠ر٠٩
٢٤	٢٢	٠	– ١ر٧	٢ر٨٩
٢٥	٢٠	١	– ٣ر٧	١٣ر٦٩
٢٦	٢٢	٠	– ١ر٧	٢ر٨٩
٢٧	١٩	١	– ٤ر٧	٢٢ر٠٩
٢٨	٢٠	٠	– ٣ر٧	٣١ر٦٩
٢٩	١٩	٠	– ٤ر٧	٢٢ر٠٩
٣٠	٢٤	٠	٠ر٣	٠ر٠٩
٣١	٢٥	٠	١ر٣	١ر٦٩
٣٢	٢٢	١	– ١ر٧	٢ر٨٩
٣٣	٢٠	٠	– ٣ر٧	١٣ر٦٩
٣٤	٢١	١	– ٢ر٧	٧ر٢٩
٣٥	٣١	٠	– ٢ر٧	٧ر٢٩
٣٦	٢٣	١	– ٠ر٧	٠ر٤٩
٣٧	٣١	٠	– ٢ر٧	٧ر٢٩
٣٨	٢٣	٠	– ٠ر٧	٠ر٤٩
٣٩	١٨	٠	– ٥ر٧	٣٢ر٤٩
٤٠	٢١	١	– ٢ر٧	٧ر٢٩

طلبة دراسات بكالوريوس

تابع / جدول (٤ - ٢)

(١) أول خانة في رقم الطالب الجامعي	(٢) العمر (س$_1$)	(٣) اختيار مادة الاحصاء ١ = نعم ٠ = لا (س$_2$)	(٤) (س$_1$ - μ)	(٥) (س$_1$ - μ)^2
٤١	١٩	٠	-٤ر٧	٢٢ر٠٩
٤٢	٢٣	٠	-٠ر٧	٠ر٤٩
٤٣	٢٢	١	-١ر٧	٢ر٨٩
٤٤	١٩	٠	-٤ر٧	٢٢ر٠٩
٤٥	٢٠	٠	-٣ر٧	١٣ر٦٩
٤٦	٢٠	٠	-٣ر٧	١٣ر٦٩
٤٧	٢١	٠	-٢ر٧	٧ر٢٩
٤٨	٢٠	١	-٣ر٧	٣١ر٦٩
٤٩	١٩	٠	-٤ر٧	٢٢ر٠٩
٥٠ ←	١٨	٠	-٥ر٧	٣٢ر٤٩
	$\sum$ س$_1$ = ١١٨٤	$\sum$ س$_2$ = ١٧	$\sum$(س$_1$ - μ) = ∴	$\sum$(س$_1$ - μ)^2 = ٨٤٤ر٩

$$\sigma^{٢} = \frac{\sum_{s=1}^{ن} (س_s - \mu)^{٢}}{ن}$$

ان العمود الرابع في الجدول (٤ - ٢) يبين الانحرافات عن الوسط، والعمود الخامس يبين مربع مجموع هذه الانحرافات . وعليه

$$\sigma^{٢} = \frac{(٢٥ - ٣٧٫٧)^{٢} + (٢٧ - ٢٣٫٧)^{٢} + \ldots + (١٨ - ٢٣٫٧)^{٢}}{٥٠}$$

$$\sigma^{٢} = \frac{(١٫٣)^{٢} + (٣٫٣)^{٢} + \ldots + (-٧٫٥)^{٢}}{٥٠}$$

$$\sigma^{٢} = \frac{١٫٦٩ + ١٠٫٨٩ + \ldots + ٣٢٫٤٩}{٥٠}$$

$$\sigma^{٢} = \frac{٨٤٤٫٩}{٥٠}$$

$$\sigma^{٢} = ٩٫١٦$$

أما بالنسبة للانحراف المعياري فهو عبارة عن الجذر التربيعي للتباين. وعليه

$$\sigma = \sqrt{٩٫١٦}$$

$$\sigma = ٤٫١$$

ـ حساب معالم مجمتع الدراسة بالنسبة للمتغيرات المتقطعة

ان العمود الثالث في الجدول (٤ ـ ٢) يعرض القيم ذات العلاقة بالمتغير المتقطع والذي هو اختيار مادة الاحصاء كمتطلب اختياري. ان هذا المتغير يعتبر متغيراً متقطعاً لان الاجابات المتوفرة عن هذا المتغير هي اما نعم أو لا. وقد تم ترقيم الاجابة نعم بالرقم (١) والاجابة لا بالرقم (٠) (صفر).

بنفس الطريقة التي تمّ بها ايجاد قيم كل من الوسط والانحراف المعياري والتباين بالنسبة للعمر، فاننا سوف نقوم بذلك بالنسبة لاختيار مادة الاحصاء.

$$\pi = \frac{\sum_{ر=1}^{ن} س_ر}{ن}$$

$$\pi = \frac{(1 + 0 + \ldots + 0)}{50}$$

$$\pi = \frac{17}{50}$$

$$\pi = ٣٤ر$$

وعليه تكون قيمة (١ ـ π) = ٦٦ر

وهذا يعني ان نسبة الأشخاص الذين اختاروا مادة الاحصاء من مجتمع الدراسة هو (٣٤ر).

أما بالنسبة للتباين فهو

$$\sigma^2 = \frac{\sum_{ر=1}^{ن} (س_ر - \pi)^2}{ن}$$

$$= \frac{(1 - ٣٤ر)^2 + (0 - ٣٤ر)^2 + \ldots + (0 - ٣٤ر)^2}{50}$$

يجب أن نلاحظ هنا ان القيمة (١ – ٣٤و)٢ تتحقق في كل مرة تكون فيها قيمة س٢ = ١ . وعليه تكون عدد المرات التي تتحقق فيها هو (٣٤و) من الحالات. وكذلك نفس الشيء بالنسبة لتحقق القيمة (٠ – ٣٤و)٢ تكون عندما تكون القيمة س٢ = صفر. وعليه تكون عدد المرات التي تحقق فيها هو (٦٦و) من الحالات.

σ^2 = (١ – ٣٤ر)^2 (٣٤ر) + (٠ – ٣٤ر)^2 (٦٦ر)

σ^2 = (٦٦ر)^2 (٣٤ر) + (– ٣٤ر)^2 (٦٦ر)

σ^2 = ٢٢٤٤ر

وعليه، فان قيمة الانحراف المعياري هي

σ = $\sqrt{\text{٢٢٤٤ر}}$

σ = ٤٧٣ر

وعليه فان المعادلة العامة لايجاد التباين للمتغيرات المتقطعة هي

$\sigma^2 = \pi\,(1 - \pi)$

= ٣٤ر (٦٦ر)

= ٢٢٤٤ر

والانحراف المعياري بالنسبة للمتغيرات المتقطعة هو

$\sigma = \sqrt{\pi\,(1 - \pi)}$

σ = $\sqrt{\text{٢٢٤٤ر}}$

σ = ٤٧٢ر

ان الذي قمنا به سابقاً هو قياس المعالم التي تصف مجتمع دراسي معلوم. دعنا الآن نقوم بعمل نفس الشيء لعينة من مجتمع دراسي. ان اهتمامنا بالنسبة للعينات هو لقياس القيم الاحصائية التي تصف العينة لعمل استنتاجات عن مدى تقدير هذه القيم الاحصاية لمعالم مجتمع الدراسة.

ـ حساب القيم الاحصائية للمتغيرات المتصلة

ان ما قمنا به سابقاً هو حساب معالم مجتمع الدراسة بالنسبة لعمر الطلاب من الجدول (٤ ـ ٢). دعنا الآن نقوم باختيار عينة من ذلك المجتمع ومن ثم نقوم بحساب الوسط والتباين والانحراف المعياري بالنسبة للعينة. ان طريقة اختيار العينة سوف تكون العينة العشوائية البسيطة.

يوجد هناك شرطان لا بد من توفرهما لتحديد وجود العينة العشوائية البسيطة. وهذان الشرطان هما :

١ ـ كل عنصر من عناصر مجتمع الدراسة له نفس فرصة الاختيار بالعينة.

٢ ـ احتمال اختيار عناصر ن عينة من مجتمع الدراسة متساوي.

لقد ذكرنا بالسابق ان اختيار العينة قد تم بالعشوائية البسيطة وكان الهدف من ذلك هو للتسهيل. والآن لا بد من توفر اجراءات أخرى والتي من خلالها يكون احتمال اختيار عناصر ن من العينات اذا كانت جميع العينات تحتوي على نفس العناصر هو متساوي. ولاختيار عناصر العينة العشوائية البسيطة فاننا سوف نستخدم جدول الأرقام العشوائية للقيام بعملية الاختيار. ان قائمة الأرقام العشوائية في الجدول الموجود في الملحق تتشكّل بطريقة من الصعب جداً التنبؤ بنمط ظهورها. فاحتمال ظهور أي رقم في هذا الجدول في أي مكان من الجدول هو نفس احتمال ظهور أي رقم. وقد تم تحديد كل طالب بمنزلتين عشريتين ما بين (٠١) و (٥٠). لذلك يمكن استخدام جدول الأرقام العشوائية للحصول على أرقام ما بين (٠١) و (٥٠) لاختيار أي عنصر من

عناصر العينة. ويمكن اختيار أرقام ذات منزلتين عشريتين من الجدول حسب العدد الذي نريده للعينة.

ولتوضيح ذلك، فاننا سوف نقوم بحل مثال. افرض أننا نريد اختيار عينة مكونة من خمسة أشخاص (ن = ٥). اننا سوف نستخدم جدول الأرقام العشوائية لذلك. ان استخدام الجدول لاختيار عينة عشوائية يتطلب تحديد الصف الذي سوف يستخدم من بين صفوف الجدول عشوائياً. وجدول الأرقام العشوائية مكون من خمسين صفاً. ثم ننتقل افقياً في الصف الذي نختاره لتحديد عناصر العينة العشوائية البسيطة. افرض أننا اخترنا الصف ٣٦ عشوائياً لاختيار عناصر العينة لذلك يجب أن نذهب الى بداية الصف ٣٦ ثم ننتقل أفقياً لاختيار عناصر العينة. فاي رقم ما بين (٠١) و (٥٠) نقوم باختياره ونقفز عن اي رقم أكبر من (٥٠). وعليه فاننا سوف نختار عناصر العينة ذات الأرقام ٣٢، ١٧، ٠٥، ٣٧، ٤١. ان هذه الأرقام هي أول خمسة أرقام ذات منزلتين عشريتين تظهر في الصف ٣٦ ما بين (٠١) و (٥٠). لذلك فان قيم العمر واختيار مادة الاحصاء في العينة سوف تكون كما يلي :

أول خانة في رقم الطالب	العمر	اختيار مادة الاحصاء
٣٢	٢٢	١
١٧	٢٦	١
٠٥	٢٥	٠
٣٧	٢١	٠
٤١	١٩	٠

وعليه فاننا نكون قد اخترنا عينة مكتوبة من خمسة أشخاص (ن = ٥) من مجتمع احصائي مكون من خمسين شخص (ن = ٥٠) . وعليه فاننا نكون قد اخترنا (ن/ن) أو (٥ / ٥٠) أو (١ر) من عناصر مجتمع الدراسة . وعليه تكون نسبة عناصر العينة الى عناصر المجتمع الدراسي هي (١ر) .

ويمكن استخدام نسبة عناصر العينة الى عناصر مجتمع الدراسة لتقدير استهلاك أو استخدام مجتمع الدراسة لسلعة أو خدمة عن طريق معرفة مجموع استخدام العينة من نفس السلعة أو الخدمة . افرض ان استهلاك العينة المكونة من (٥) أشخاص من مادة البيض هو (٣٥) بيضة أسبوعياً، فان حساب استهلاك مجتمع الدراسة من مادة البيض سوف يكون كما يلي :

$$\text{استخدام مجتمع الدراسة} = \frac{\text{استخدام العينة من مادة البيض}}{\text{نسبة العينة}}$$

$$= \frac{٢٥}{١ر}$$

$$= ٣٥٠ \text{ بيضة} .$$

ان الوسط أو المعدل للعينة يساوي مجموع القيم مقسوماً على عدد افراد العينة أو حجم العينة . وعليه فان الوسط

$$\bar{س} = \frac{\sum_{ر=١}^{ن} س_ر}{ن}$$

$$= \frac{١٩ + ٢١ + ٢٥ + ٢٦ + ٢٢}{٥}$$

$$= \frac{١١٣}{٥}$$

$$= ٢٢٫٦$$

وهذا يعني ان قيمة الوسط للعينة هو (٢٢ر٦). اننا نلاحظ ان قيمة الوسط الحسابي للعينة أقل بشيء بسيط عن الوسط الحسابي الحقيقي لعمر المجتمع والذي هو (٢٣ر٧). ان الوسط لمجتمع الدراسة لمعظم المشاكل الحقيقية يكون غير معروف. لذلك فاننا سوف نستخدم قيمة الوسط الحسابي للعينة كأفضل تقدير لقيمة الوسط الحقيقي وهو قيمة الوسط لمجتمع الدراسة.

أما بالنسبة للتباين للعينة فهو يساوي مجموع مربع الانحرافات عن وسطها الحسابي مقسوماً على درجات الحرية. وعليه يكون التباين للعينة

$$S^2 = \frac{\sum_{ر=١}^{ن} (س_ر - \bar{س})^2}{\text{درجات الحرية}}$$

درجات الحرية = حجم العينة – عدد التقديرات (عدد القيم الاحصائية المحسوبة).

ان قيمة الوسط للعينة هي القيمة المحسوبة فقط. بمعنى لقد تم تقدير قيمة الوسط الحقيقي لمجتمع الدراسة بقيمة الوسط المحسوب من العينة. وهذا التقدير يقلّل درجات الحرية بدرجة واحدة. وعليه تكون

درجات الحرية = ٥ – ١ = ٤

$$S^2 = \frac{(٢٢ - ٢٢ر٦)^2 + (٢٦ - ٢٢ر٦)^2 + (٢٥ - ٢٢ر٦)^2 + (١٩ - ٢٢ر٦)^2}{٤}$$

$$= \frac{[(-٦ر)^2 + (٣ر٤)^2 + (٢ر٤)^2 + (-١ر٦)^2 + (-٣ر٦)^2]}{٤}$$

$$= \frac{(٣٦ر + ١١ر٥٦ + ٥ر٧٦ + ٢ر٥٦ + ١٢ر٩٦)}{٤}$$

$$= \frac{٣٣٫٢}{٤}$$

$$= ٨٫٣$$

لقد كان التباين لمجتمع الدراسة يساوي (١٦٫٩)، وعليه نلاحظ استخدام العينة عمل على تقليل التباين. كما ان هذا الفرق يزداد لو قمنا بالقسمة على ن بدلاً من ن – ١ .

وعليه فان الانحراف المعياري يساوي

$$S = \sqrt{٨٫٣}$$

$$S = ٢٫٨٨$$

ولتسهيل العملية الحسابية في ايجاد التباين، فانه يمكن استخدام المعادلة التالية

$$S^2 = \frac{\sum س_ر^2 - \left[\frac{(\sum س_ر)^2}{ن}\right]}{ن - ١}$$

والانحراف المعياري

$$S = \sqrt{\frac{\sum س_ر^2 - \left[\frac{(\sum س_ر)^2}{ن}\right]}{ن - ١}}$$

ولحل المثال السابق باستخدام هذه المعادلة، فاننا نقوم بايجاد ما يلي :

$$\sum س_ر^2 = (٢٢)^2 + (٢٦)^2 + (٢٥)^2 + (٢١)^2 + (١٩)^2 = ٢٥٨٧$$

$$(\sum س_ر)^2 = (٢٢ + ٢٦ + ٢٥ + ٢١ + ١٩) = ١٢٧٦٩$$

وعليه :

$$S^{2} = \frac{٢٥٨٧ - \frac{١٢٧٦٩}{٥}}{٤}$$

$$= \frac{٢٥٨٧ - ٢٥٥٣ر٨}{٤}$$

$$= \frac{٣٣ر٢}{٤}$$

$$= ٨ر٣$$

اما بالنسبة للانحراف المعياري

$$S = \sqrt{٨ر٣}$$

$$= ٢ر٨٨$$

نلاحظ ان قيم التباين والانحراف المعياري باستخدام هذه المعادلة هو نفس قيم التباين والانحراف المعياري بالمعادلة السابقة. وعليه فان جميع الحسابات لـ S^{2} و S سوف يكون بالمعادلة الثانية لما توفره من وقت كثير في الحسابات.

- حساب القيم الاحصائية للمتغيرات المتقطعة.

دعنا الآن نقوم بحساب الوسط والتباين والانحراف المعياري للمتغير المتقطع وهو اختيار مادة الاحصاء.

$$\rho = \frac{\sum_{\text{ر}=١}^{\text{ن}} \text{س}_{\text{ر}}}{\text{ن}}$$

$$\rho = \frac{١ + ١ + ٠ + ٠ + ٠}{٥}$$

$$\rho = \frac{٢}{٥}$$

$$\rho = \text{٤ر}$$

بينما $(١ - \rho) = (١ - \text{٤ر}) = \text{٦ر}$

أما بالنسبة للتباين

$$S^{٢} = \frac{\sum (\text{س}_{\text{ر}} - \rho)^{٢}}{\text{ن} - ١}$$

$$S^{٢} = \frac{(١ - \text{٤ر})^{٢} + (١ - \text{٤ر})^{٢} + (٠ - \text{٤ر})^{٢} + (٠ - \text{٤ر})^{٢} + (٠ - \text{٤ر})^{٢}}{٤}$$

$$S^{٢} = \frac{(\text{٦ر})^{٢} + (\text{٦ر})^{٢} + (-\text{٤ر})^{٢} + (-\text{٤ر})^{٢} + (-\text{٤ر})^{٢}}{٤}$$

$$S^{٢} = \frac{\text{١ر٢}}{٤}$$

$$S^{٢} = \text{٣ر}$$

أما بالنسبة للانحراف المعياري

$$S = \sqrt{\text{٣ر}}$$

$$S = \text{٥٤٨ر}$$

نلاحظ ان المعادلة العامة لحساب التباين هي

$$S^{2} = \frac{\text{ن} \times p \times (\text{١} - p)}{\text{ن} - \text{١}} \qquad q = (\text{١} - p)$$

$$S^{2} = \frac{\text{ن} \times p \times q}{\text{ن} - \text{١}}$$

$$S^{2} = \frac{(\text{٥})(\text{٤ر})(\text{٦ر})}{\text{٤}}$$

$$S^{2} = \text{٣ر}$$

أما بالنسبة للانحراف المعياري

$$S = \sqrt{\frac{\text{ن} \times p \times q}{\text{ن} - \text{١}}}$$

ان هذه القيم الاحصائية المحسوبة هي عبارة عن قيم تقديرية لمعالم مجتمع الدراسة.

- عمل الاستنتاجات عن معالم مجتمع الدراسة من القيم الاحصائية للمتغيرات المتصلة.

- النظرية الأساسية

ان الذي قمنا بحسابه في السابق هو وسط العمر للعينة التي تمّ اختيارها عشوائياً من مجتمع الطلبة. وحتى نتمكن من معرفة مدى قدرة القيمة الاحصائية في تقدير معلمة المجتمع الاحصائي فانه لا بد من فهم نظرية الاستنتاجات الاحصائية.

دعنا أولاً نفترض اننا قمنا بسحب عينة عشوائية أخرى من مجتمع الطلبة تتكون من خمسة أشخاص أيضاً. وقد قمنا بتسمية أفراد العينة الجديدة بالأحرف أ، ب، جـ، د، هـ. ونريد الآن اختيار عينة عشوائية واحدة تتكون من فردين (ن = ٢) من هذا المجتمع. لاختيار هذه العينة فانه يوجد هناك عدداً من التوافيق بالنسبة للعناصر التي يمكن ان تشكل العينة. بمعنى آخر، يوجد هناك عدداً من العينات المختلفة وهما

رقم العينة	عناصر العينة
١	أ ب
٢	أ جـ
٣	أ د
٤	أ هـ
٥	ب جـ
٦	ب د
٧	ب هـ
٨	جـ د
٩	جـ هـ
١٠	د هـ

نلاحظ وجود (١٠) توافيق ممكنة لاختيار عنصرين من مجتمع الدراسة الذي يتكون من (٥) عناصر. ويمكن حساب ذلك رياضياً عن طريق تحديد عدد التوافيق الممكنة باستخدام المعادلة التالية :

$$\text{توافيق } {}^{ن}_{ن} = \binom{ن}{ن} = \text{توافيق } (ن ، ن)$$

$$= \frac{ن!}{ن(ن - ن)!}$$

ولحل المثال السابق، نلاحظ ان ن = ٥ ، ن = ٢ ، لذلك

$$\text{توافيق } (٥ ، ٢) = \frac{٥!}{٢! \ ٣!}$$

$$= \frac{٥ \times ٤ \times ٣ \times ٢ \times ١}{٢ \times ١ \times ٣ \times ٢ \times ١}$$

$$= \frac{٢٠}{٢}$$

$$= ١٠$$

فكل عينة من هذه العينات العشرة ممكن ان ينتج عنها تقدير لوسط مجتمع الدراسة.

دعنا الآن نعود الى مثالنا الأصلي بالنسبة لمجتمع الدراسة الذي يتكون من (٥٠) طالب كما يوضح الجدول (٤ - ٢). لقد قمنا بالسابق باختيار عينة واحدة من مجتمع الدراسة تتكون من خمسة طلاب. وعليه فان عدد العينات المختلفة التي يمكن اختيارها

من مجتمع الدراسة وتتكون من خمسة افراد هو

$$\text{توافيق } (٥٠ ، ٥) = \frac{٥٠!}{٥! \; ٤٥!}$$

$$= \frac{٥٠ \times ٤٩ \times ٤٨ \times ٤٧ \times ٤٦ \times ٤٥!}{٥ \times ٤ \times ٣ \times ٢ \times ١ \times ٤٥!}$$

$$= \frac{٢٥٤٢٥١٢٠٠}{١٢٠}$$

$$= ٢١١٨٧٦٠ \text{ عينة ممكنة}$$

ونحن قمنا فقط باختيار عينة واحدة فقط من عدد العينات الممكن (٢١١٨٧٦٠) .

ان النقطة المهمة التي يجب ادراكها هو عدد العينات الممكن اختيارها من أي مجتمع دراسة . ويجب أن نفهم هنا ان الاستنتاج الاحصائي التقليدي يعتمد على ما سيحدث اذا تم اختيار أكثر من عينة مختلفة من مجتمع الدراسة .

لقد قمنا بالسابق اختيار عينة عشوائية واحدة تتكون من خمسة افراد وكان وسط أعمار أفراد العينة هو (٢٢ر٦) . افرض اننا قمنا باختيار عينة عشوائية أخرى تتكون أيضاً من خمسة طلاب وتم حساب الوسط لأعمار الطلبة في هذه العينة وكان (٢٣ر٤) . افرض أيضاً ان قمنا باختيار عينة ثالثة تتكون من خمسة طلاب وتم حساب وسط أعمار أفراد العينة الثالثة، فانه من الممكن الحصول على وسط قدره (٢٤ر٢) . ان القيم المختلفة لأوساط أعمار الطلبة للعينات المختلفة ممكن أن يؤدي الى الارباك لأن هذا يعني وجود ثلاثة تقديرات مختلفة لمعلم مجتمع الدراسة .

ان النظرية الاحصائية لا تتوقف عند اختيار فقط ثلاثة عينات، بل تقوم على

اختيار عينات مختلفة أيضاً وأيضاً. ان الاختيار المتكرر للعينات قد يؤدي الى تكرار بعض القيم الاحصائية لنفسها. وبالتحديد فاننا سوف نلاحظ ان القيمة الاحصائية القريبة من القيمة الحقيقية لمجتمع الدراسة تعمل على تكرار نفسها أكثر من القيمة الاحصائية البعيدة عن القيمة الحقيقية لمجتمع الدراسة. فاذا ما قمنا بتمثيل هذه القيم المختلفة للوسط، فان ذلك قد يؤدي الى تشكيل منحنى جرسي (طبيعي). ويطلق على توزيع أوساط العينة بالتوزيع العيني للوسط أو باختصار التوزيع العيني. ويعتبر التوزيع العيني مهم لسببين.

١ - أوساط العينة في هذا التوزيع تكون موزعة حول الوسط الحقيقي لمجتمع الدراسة بطريقة معروفة.

٢ - نستطيع باستخدام هذا التوزيع من تحديد مدى قرب توزيع القيم الاحصائية عن معالم مجتمع الدراسة.

وحتى نتمكن من الاستفادة من التوزيع العيني فانه لا بد من تحديد طبيعة التوزيع العيني للوسط. وللقيام بذلك فانه لا بد من الرجوع الى نظرية الحد المركزية في علم الاحصاء والتي تقول :

١ - اذا كان التوزيع لمجتمع الدراسة لقياس معين طبيعي، فان التوزيع العيني للوسط لجميع العينات سوف يكون توزيعاً طبيعياً أيضاً.

٢ - اذا كان التوزيع لمجتمع الدراسة لقياس ليس طبيعياً، فان التوزيع العيني للوسط يقترب من التوزيع الطبيعي كلما زاد عدد أفراد العينة.

٣ - ان وسط التوزيع العيني للوسط يساوي وسط مجتمع الدراسة. وفي الحالة التي تكون فيها القيمة المتوقعة للوسط مساوية لقيمة معلمة مجتمع الدراسة، فاننا نقول بان القيمة الاحصائية غير متحيزة.

٤ - ان الانحراف المعياري للتوزيع العيني للوسط هو الانحراف المعياري بالنسبة لمجتمع

الدراسة مقسوماً على الجذر التربيعي لحجم العينة. وعليه

$$\sigma \bar{x} = \frac{\sigma}{\sqrt{ن}}$$

وغالباً ما يطلق على هذه القيمة بالخطأ المعياري للوسط.

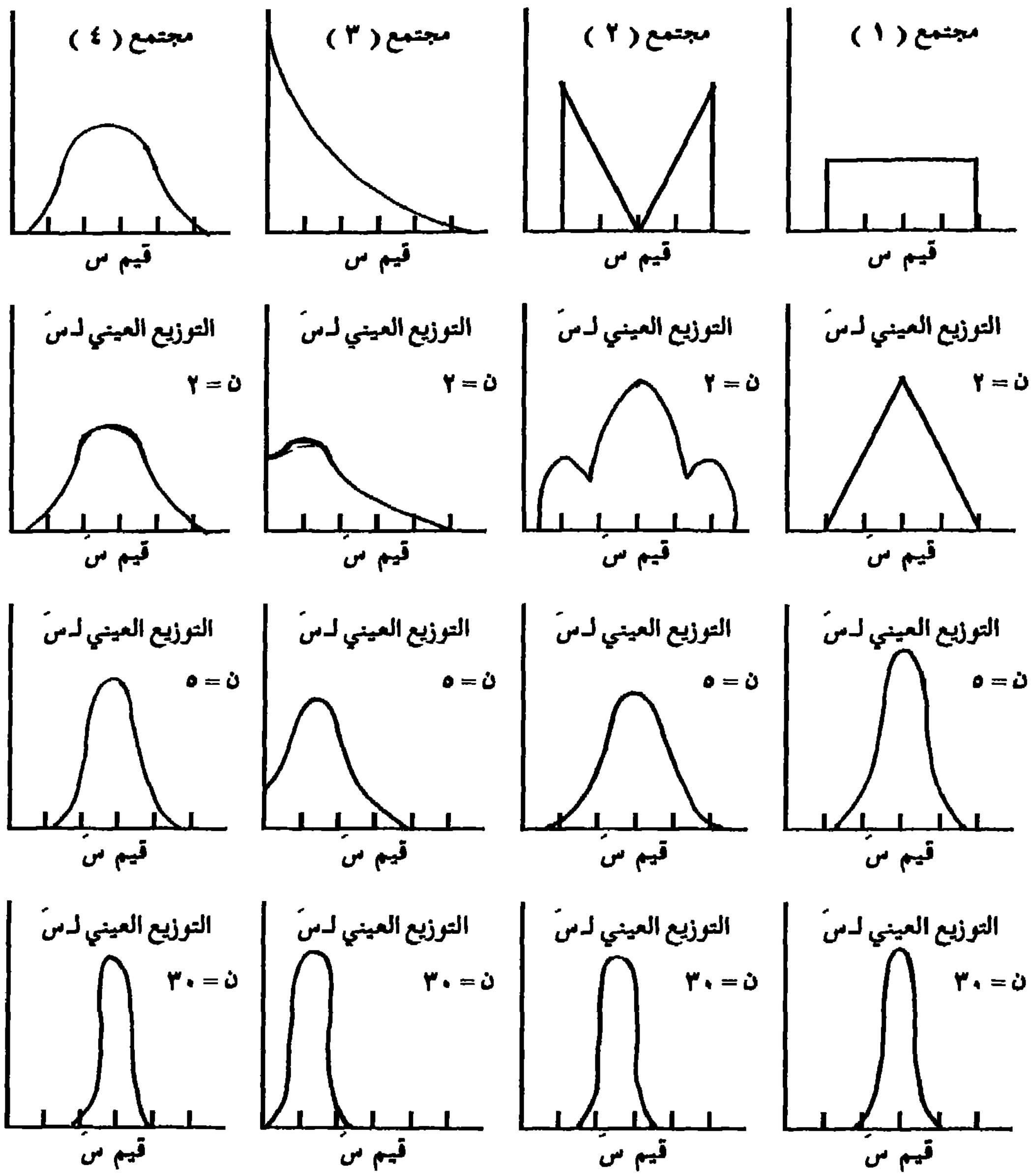
ولا نستطيع بالحياة العملية من معرفة قيمة كل من μ أو σ . لذلك فاننا سوف نقوم بتقديرهم بالقيم الاحصائية سَ و S على التوالي والتي تحسب قيمهم بناءاً على عينة تم اختيارها من مجتمع الدراسة. لذلك

$$S \bar{x} = \frac{S}{\sqrt{ن}}$$

ويوضح الشكل (٤ - ١) نظرية الحد المركزية. ويظهر في الشكل توزيعات لأربعة مجتمعات مختلفة أحدهما موزعاً توزيعاً طبيعياً وهو المجتمع الذي يقع في اليسار. ولكن نلاحظ ان التوزيع العيني للوسط لكل مجتمع من هذه المجتمعات يكون بالنهاية توزيعاً طبيعياً عندما يكون عدد أفراد العينة يساوي (٥). أما بالنسبة للتوزيع العيني للوسط الذي تم اختياره من مجتمع طبيعي هو توزيع طبيعي كامل. كما نلاحظ ان التوزيع العيني للوسط لجميع العينات هو توزيع طبيعي عند يصبح حجم العينة يساوي (٣٠) أو أكثر. لذلك عندما نقوم بحساب الوسط لأي عينة يكون حجمها (٣٠) فاننا نكون متأكدون بأن الوسط يأتي من توزيع طبيعي وسطه يساوى μ وخطأه المعياري يساوي $\sigma / \sqrt{ن}$.

كما ان هناك خاصية أخرى يجب أن تكون مفهومة عن المنحى الطبيعي قبل استخدام هذه النظرية. وتتعلق هذه الخاصية بالمساحة المحصورة تحت المنحى الطبيعي. والشكل (٤ - ٢) يوضح منحى طبيعي بمساحات مختلفة واقعة ما بين انحرافات معيارية مختلفة عن الوسط.

شكل (٤ ـ ٢)



Source : Earmest K. , Gerald J. , and Fredrick R. , Statistics for Business Decisions (Homewood, 111, Irwin, 1959, pp. 181 - 183) .

وبالتحديد يمكن ملاحظة ما يلي تقريباً :

١ - ٦٨٪ من الحالات تقع ما بين ± ١ انحراف معياري عن الوسط.

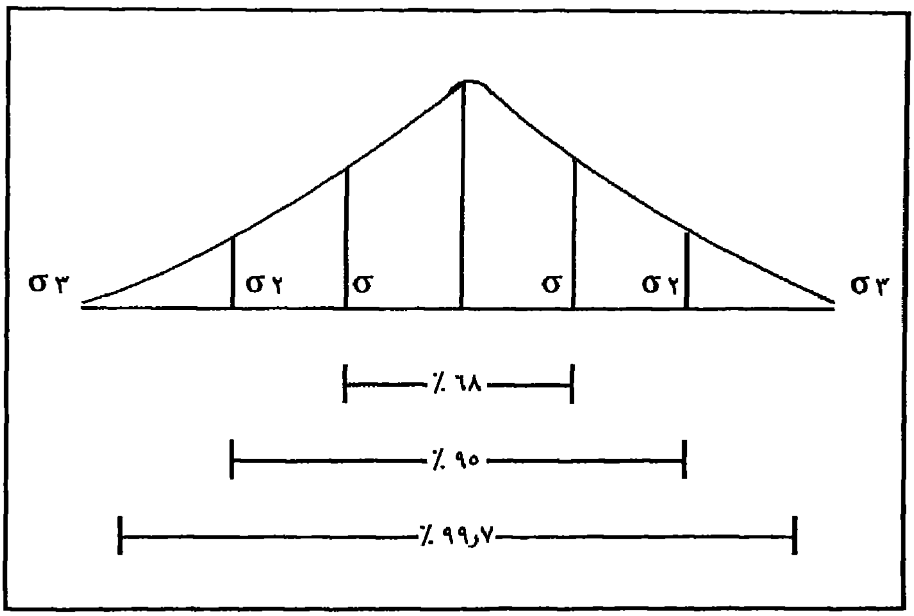
٢ - ٩٥٪ من الحالات تقع ما بين ± ٢ انحراف معياري عن الوسط.

٣ - ٩٩ر٧٪ من الحالات تقع ما بين ± ٣ انحراف معياري عن الوسط.

اننا الآن على علم بكل شيء نحتاجه عن النظرية لنتمكن من تحديد كم يكون متوسط العينة جيداً في تقدير متوسط المجتمع.

شكل (٤ - ٢)

المساحة تحت المنحنى الطبيعي



ـ استخدام النظرية

لقد تمّ اختيار العينة المكونة من اعمار خمسة طلاب من مجتمع كان توزيعه قريباً من التوزيع الطبيعي. لذلك فان الوسط الحسابي الذي قمنا بحسابه جاء من التوزيع الطبيعي للأوساط. وقد قمنا بتقدير ما يلي :

١ ـ وسط التوزيع العيني بـ $\bar{\text{س}}$

٢ ـ الانحراف المعياري للتوزيع العيني (الخطأ المعياري) بـ $S\bar{x}$.

$$S\bar{x} = \frac{S}{\sqrt{\text{ن}}}$$

$$= \sqrt{\frac{\sum_{\text{ر}=١}^{\text{ن}} (\text{س}_{\text{ر}} - \bar{\text{س}})^{٢} / (\text{ن} - ١)}{\sqrt{\text{ن} - ١}}}$$

وفي مثالنا، كانت $\bar{X}$ = ٢٢ر٦ و S = ٢ر٨٨، وعليه

$$S\bar{x} = \frac{٢ر٨٨}{\sqrt{٥}} = \frac{٢ر٨٨}{٢ر٢٤}$$

$$S\bar{x} = ١ر٣$$

دعنا الآن نقوم بحساب حجم الفترات عندما يكون بعد الانحراف المعياري عن الوسط ± ١ ثم ± ٢ ثم ± ٣ انحرافات معيارية عن الوسط.

عند ± ١ انحراف معياري، فان الفترة تكون

٢٢ر٥ ± ١ر٣ = ٢١ر٣ – ٢٣ر٩

اننا نعلم بان ٦٨٪ من أوساط التوزيع العيني يقع في هذه الفترة اذا ما كان وسط

العينة $\bar{X}$ المحسوب هو فعلاً وسط التوزيع العيني. (نذكر ان μ = ٧ر٢٣).

أما بالنسبة للفترة عند ± ٢ انحرافين معيارين فهي

٦ر٢٢ ±٢ (٣ر١) = ٦ر٢٢ ± ٦ر٢

= ٢٠ – ٢ر٢٥

اننا نعلم بان ٩٥٪ من أوساط التوزيع العيني يقع في هذه الفترة اذا ما كان وسط العينة المحسوب $\bar{X}$ هو فعلاً وسط التوزيع العيني.

أما بالنسبة للفترة ± ٣ انحرافات معيارية فهي

٦ر٢٢ ±٣ (٣ر١) = ٦ر٢٢ ± ٩ر٣

= ٧ر١٨ – ٥ر٢٦

اننا نعلم بان ٧ر٩٩٪ من أوساط التوزيع العيني يقع ضمن هذه الفترة اذا ما كان وسط العينة المحسوب $\bar{X}$ فعلاً وسط التوزيع العيني.

اننا ندعوا الفترات التي تمّ حسابها بفترات الثقة. وقد كانت فترة الثقة الأولى هي فترة الثقة الـ ٦٨٪، والثانية فترة ثقة الـ ٩٥٪ والثالثة كانت فترة ثقة الـ ٧ر٩٩٪. لاحظ انه يجب تحديد مستوى الثقة قبل فترة الثقة. وحتى نتمكن من حساب فترة الثقة فانه لا بد من أن يأتي الوسط من توزيع طبيعي للأوساط.

ولمعرفة ما تعنيه فترة الثقة وبدقة، دعنا نستخدم فترة ثقة الـ ٩٥٪ للتوضيح. اننا نعرف الآن مما سبق بان وسط التوزيع العيني هو وسط المجتمع. واننا نستخدم وسط عينة واحدة ($\bar{X}$) لتقدير وسط التوزيع العيني. كما اننا قد قمنا باختيار عينة واحدة فقط. لذلك فاننا لا نستطيع تحديد ما هو وسط التوزيع العيني الصحيح. وعليه فان من الممكن اختيار عينة بحيث لا يكون وسط المجتمع الحقيقي ضمن فترة ثقة الـ ٩٥٪ المحسوبة من العينة. لهذا فالوسط بالنسبة لمجتمع الدراسة قد يكون أو قد لا يكون ضمن فترة ثقة الـ ٩٥٪ المحسوبة من العينة. لذلك فان فترة ثقة الـ ٩٥٪ لا تعني ان احتمال

ظهور وسط المجتمع في فترة الثقة هو ٩٥٪ . ولكن ما تعنيه فترة ثقة الـ ٩٥٪ هو اذا تمّ اختيار ١٠٠ عينة عشوائية مختلفة ومن ثم قمنا بحساب (١٠٠) فترة ثقة ٩٥٪، فان نتوقع ان يظهر وسط المجتمع في ٩٥ فترة ثقة من الـ (١٠٠) فترة.

لقد كان حدود فترة ثقة الـ ٩٥٪ المحسوبة هي من ٢٠ الى ٢ر٢٥ . واذا ما تم اختيار (١٠٠) عينة مختلفة، فان وسط المجتمع سوف يظهر في فترات ٩٥ عينة من الـ (١٠٠) عينة. وفي مثالنا نلاحظ ان وسط المجتمع (μ = ٧ر٢٣) يقع ضمن حدود فترة ثقة الـ ٩٥٪ التي تم حسابها من العينة التي حجمها يساوي خمسة. لذلك يمكن القول بان دقة التقدير عند مستوى الثقة ٩٥٪ هو ± ٦ر٢ سنة. لاحظ ان حجم الدقة له معنى فقط عند مستوى ثقة معينة. فعند مستوى الثقة ٧ر٩٩٪ تكون دقة التقدير هي ± ٩ر٣ سنة. وهذا يعني ان حجم الدقة يساوي عرض فترة الثقة.

ان حساب فترة الثقة يزودنا بقياس خطأ العينة ودقة العينة المختارة. ويجب التنويه هنا الى ان اجراءات الاختيار الاحتمالية للعينة هي التي تساعد في حساب فترة ثقة مناسبة. بما أن معرفتنا عن التوزيع العيني للأوساط من اجراءات الاختيار اللااحتمالية للعينة معدومة فاننا لا نستطيع حساب فترة ثقة ذات معنى.

ـ ما هو تأثير حجم العينة على الدقة.

لقد تم مسبقاً تحديد المعادلة المستخدمة في ايجاد الانحراف المعياري للتوزيع العيني للوسط. وهذه المعادلة هي :

$$S\bar{x} = \sqrt{\frac{\sum (س_ر - \bar{س})^2 / (ن - ١)}{\sqrt{ن}}}$$

وببساطة اكثر

$$S\bar{x} = \frac{S}{\sqrt{ن}}$$

علماً بان (S) هو الانحراف المعياري للتوزيع المتغير ذي الاهتمام. اننا نلاحظ ان قيمة $S\bar{x}$ تتغير عكسياً مع الجذر التربيعي لحجم العينة التي نختارها. بمعنى آخر، كلما زاد حجم العينة كلما زادت قيمة الجذر التربيعي لحجم العينة وعليه كلما قلّت قيمة $S\bar{x}$. لاحظ ان فترة الثقة حول وسط العينة $\bar{x}$ سوف تتناقص كلما قلّت قيمة $S\bar{x}$ ، وعليه كلما زاد التأكد بالنسبة لدقة التقدير.

ويمكن توضيح ذلك من خلال مثال مجتمع الطلبة. ولكن افرض هذه المرة ان عدد عناصر المجتمع هو (٥٠٠٫٠٠٠) بدلاً من ٥٠، وان وسط المجتمع والانحراف المعياري كما كان. افرض اننا قمنا باختيار عدد من العينات العشوائية البسيطة من مجتمع الدراسة، وقد كان الوسط والتباين المحسوب لجميع العينات متماثل وكما يلي

$\bar{x}$ = ٢٢٫٦ ، S^2 = ٨٫٣ ، S = ٢٫٨٨

بينما كان حجم العينات المختلفة التي تم اختيارها كما يلي :

١ ـ ن = ٥ ٢ ـ ن = ٣٠ ٣ـ ن = ١٠٠

٤ ـ ن = ١٠٠٠ ٥ ـ ن = ٢٠٠٠

دعنا الآن نقوم بحساب $S\bar{x}$ لكل عينة

١ ـ $S\bar{x} = \frac{٢٫٨٨}{\sqrt{٥}} = ١٫٣$

٢ ـ $S\bar{x} = \frac{٢٫٨٨}{\sqrt{٣٠}} = \frac{٢٫٨٨}{٥٫٤٨} = ٫٥٣$

٣ ـ $S\bar{x} = \frac{٢٫٨٨}{\sqrt{١٠٠}} = \frac{٢٫٨٨}{١٠} = ٫٢٩$

٤ - $S_{\bar{x}}$ = ٢ر٨٨ / √١٠٠٠ = ٢ر٨٨ / ٣١ر٧ = ر٠٩

٥ - $S_{\bar{x}}$ = ٢ر٨٨ / √٢٠٠٠ = ٢ر٨٨ / ٤٤ر٨ = ر٠٦

دعنا نحسب الآن فترة الثقة الـ ٩٥٪ لكل عينة وبالتسلسل.

(١) ٢٢ر٦ ± ٢ (١ر٣) = ٢٢ر٦ ± ٢ر٦ = ٢٠ – ٢٥ر٢

(٢) ٢٢ر٦ ± ٢ (ر٥٣) = ٢٢ر٦ ± ١ر٠٦ = ٢١ر٥ – ٢٣ر٧

(٣) ٢٢ر٦ ± ٢ (ر٢٩) = ٢٢ر٦ ± ر٥٨ = ٢٢ر٠٢ – ٢٣ر١٨

(٤) ٢٢ر٦ ± ٢ (ر٠٩) = ٢٢ر٦ ± ر١٨ = ٢٢ر٤٢ – ٢٢ر٧٨

(٥) ٢٢ر٦ ± ٢ (ر٠٦) = ٢٢ر٦ ± ر١٢ = ٢٢ر٤٨ – ٢٢ر٧٢

نلاحظ انه عندما زاد حجم العينة ن من ٥ الى ٢٠٠٠، نجد ان طول فترة الثقة الـ ٩٥٪ تتناقص من (٢ر٥) الى (٢ر).

وهذا يعني ان الثقة في التقدير تزداد مع تزايد حجم العينة. بمعنى آخر اذا كان حجم العينة ن = ٢٠٠٠ فاننا نتوقع ان يقع الوسط الحقيقي في فترة ثقة طولها فقط ٢ر في ٩٥ عينة من أصل ١٠٠ عينة.

وفي مثالنا نلاحظ ان الوسط الحقيقي (μ = ٢٣ر٧) لا يقع ضمن فترة الثقة الـ ٩٥٪ والتي حدودها من ٢٢ر٤٨ – ٢٢ر٧٢. وقد حدث هذا لافتراضنا ان أوساط العينة متساوية وبغض النظر عن حجم العينة. وقد قمنا بذلك لتوضيح تأثير حجم العينة على طول فترة الثقة المحسوبة. وبالحقيقة فانه كلما زاد حجم العينة كلما اقتربنا من القيمة الحقيقية لوسط المجتمع. والقيمة الاحصائية التي تقترب من قيمة معلمة المجتمع كلما زاد حجم العينة تدعى تقدير مستمر. وبالتأكيد نحن نبحث دائماً عن التقادير المستمر.

لذلك نتوقع أن يكون وسط العينة التي حجمها يساوي (٢٠٠٠) أقرب الى وسط المجتمع (μ = ٧ر٢٣) مقارنة مع العينة التي حجمها يساوي (٥). بمعنى آخر، كلما زاد عدد أفراد العينة كلما زادت درجة تمثيل العينة لمجتمع الدراسة.

ـ عمل استنتاجات عن معالم مجتمع الدراسة من عينة احصائية لمتغيرات متقطعة.

ان كل النظرية والاجراءات التي تم تطبيقها على المتغيرات المتصلة يمكن تطبيقها على المتغيرات المتقطعة. ان وسط العينة لنسبة الطلبة الذين اختاروا مادة الاحصاء في العينة التي تم اختيارها وكان عدد أفرادها خمسة أشخاص هو (٤ر). ان هذا الوسط هو وسط واحد فقط من توزيع الأوساط للنسب الذي ينتج عن تكرار اختيار عينات متعددة حيث ان عدد أفراد كل منها هو خمسة أشخاص. ويمكن هنا ايضاً تطبيق نظرية الحد المركزية. وفي هذه الحالة يكون وسط التوزيع العيني للأوساط هو π والانحراف المعياري للتوزيع العيني هو $\sqrt{\pi(١-\pi)/N}$. ونحن لا نعلم القيم الحقيقية لمجتمع الدراسة بالحياة العملية. لذلك فاننا نقوم بتقديرهم كما يلي :

الوسط = p

الخطأ المعياري = $S_p = \sqrt{\frac{p(١-p)}{ن}}$ ، $q = (١-p)$

وسوف نقوم الآن بحساب فترة ثقة ٩٥٪ للعينة قيد الدراسة والتي عدد عناصرها يساوي (٥). ان المعادلة المستخدمة هي

$$p \pm ٢\sqrt{\frac{pq}{ن}}$$

وفي مثالنا كانت قيمة p = ٤ر ، ن = ٥ . وعليه تكون فترة ثقة الـ ٩٥٪ هي

$$\text{٤ر} \pm \text{٢} \sqrt{\frac{(\text{٤ر})(\text{٦ر})}{\text{٥}}}$$

$$= \text{٤ر} \pm \text{٢} \sqrt{\text{٠٤٨ر}}$$

$$= \text{٤ر} \pm \text{٢} (\text{٢٢ر})$$

$$= \text{٤ر} \pm \text{٤٤ر}$$

$$= \text{٠} - \text{٨٤ر}$$

ان هذه الفترة عريضة جداً وذلك لان حجم العينة قليل وكذلك قيمة qp للمتغير هي كبيرة. إن ما يمكن قوله هنا اننا نتوقع وجود وسط النسبة الحقيقي لمجتمع الدراسة في هذا النوع من الفئات في ٩٥ عينة من أصل (١٠٠) عينة. ان الوسط الحقيقي لمجتمع الدراسة (π = ٢٧ر) هو موجود في هذه الفترة.

دعنا الآن نلاحظ تاثير زيادة حجم العينة الى (٢٠٠٠) على فترة الثقة. اذا كانت p = ٤ر ، فان فترة ثقة ٩٥٪ سوف تكون

$$\text{٤ر} \pm \text{٢} \sqrt{\frac{(\text{٤ر})(\text{٦ر})}{\text{٥}}}$$

$$= \text{٤ر} \pm \text{٢} \sqrt{\text{٠٠٠١٢ر}}$$

$$= \text{٤ر} \pm \text{٢} \sqrt{\text{٠١١ر}}$$

$$= \text{٤ر} \pm \text{٠٢٢ر}$$

$$= \text{٣٨ر} - \text{٤٢ر}$$

ان سبب وجود الوسط الحقيقي لمجتمع الدراسة خارج حدود فترة الثقة هو تثبيت قيمة p = ٤و والتي حصلنا عليها من العينة التي كان عدد عناصرها هو خمسة أشخاص وليس (٢٠٠٠). فاذا كان عدد أفراد العينة يساوي (٢٠٠٠) فاننا نتوقع أن يقترب

الوسط أكثر الى القيمة الحقيقية لمجتمع الدراسة مقارنة مع الوسط الذي نحصل عليه من العينة التي عدد افرادها كان (٥) .

لاحظ ان أعلى قيمة لـ $S\rho$ هو عندما تكون قيمة ρ = (٥ر) وبغض النظر عن حجم العينة لأن $q\rho$ = ٢٥ر . ولا يمكن لـ $q\rho$ أن يأخذ قيمة أعلى من هذه القيمة . فاذا كانت فيه قيمة ρ = ٦ر ، فان $q\rho$ = ٢٤ر . اذا كانت قيمة ρ = ٧ر ، فان قيمة $q\rho$ = ٢١ر ، اذا كانت قيمة ρ = ٨ر ، فان قيمة $q\rho$ = ١٦ر . وهذا يعني انه كلما زادت عدم دقة النتائج (كلما اقتربت فيه ρ من ٥ر) كلما زادت قيمة الخطأ .

وعليه فقد قمنا الآن بحساب قيمة الخطأ العيني للمتغيرات المتصلة والمتقطعة .

ـ السؤال بالنسبة لحجم المجتمع .

لم يذكر الشرح لحد الآن ي شيء عن أهمية حجم مجتمع الدراسة في الحسابات التي تم ايجادها . ان هذا من حيث المبدأ أمر غير صحيح . ولكن بما أن معظم مجتمعات الدراسة بالنسبة للأبحاث الادارية كبيرة فان هذا يجعل الاهتمام بحجم مجتمع الدراسة أمر غير مهم .

ان مشكلة حجم الدراسة غالباً ما تظهر في المجتمعات المحدودة فقط . لذلك لا بد من تغيير المعادلة المستخدمة في ايجاد الخطأ المعياري للتوزيع العيني للوسط بالنسبة للمجتمعات المحدودة . وللتخلص من هذه المشكلة فاننا نقوم باضافة معامل التصحيح المحدد للمعادلة السابقة وكما يلي :

(١) أ) $$\sigma\bar{x} = \frac{\sigma x}{\sqrt{ن}} \sqrt{\frac{ن' - ن}{ن' - ١}}$$ ، حيث ن' = حجم المجتمع

ن = حجم العينة

ب) $$S\bar{x} = \frac{Sx}{\sqrt{ن}} \sqrt{\frac{ن' - ن}{ن' - ١}}$$

$$(٢)\ (أ \quad \sigma\rho = \sqrt{\frac{\pi(1-\pi)}{\text{ن}}}\ \sqrt{\frac{\text{نَ} - \text{ن}}{\text{نَ} - ١}}$$

$$(ب \quad S\rho = \sqrt{\frac{q\rho}{\text{ن}}}\ \sqrt{\frac{\text{نَ} - \text{ن}}{\text{نَ} - ١}}$$

ان قيمة $\sqrt{(\text{نَ} - \text{ر}) / (\text{نَ} - ١)}$ تدعى بقيمة معامل التصحيح المحدّد. فاذا كان حجم مجتمع الدراسة (نَ) كبير بالنسبة لحجم العينة (ن) ، فان قيمة $\sqrt{(\text{نَ} - \text{ر}) / (\text{نَ} - ١)}$ تساوي تقريباً (١). لذلك فاننا نستطيع استخدام المعادلة الأساسية لـ $\sigma\bar{X}$ و $S\bar{X}$ وكذلك نفس الشيء بالنسبة $\sigma\rho$ و $S\rho$.

ان مدى معامل التصحيح يقع دائماً ما بين الصفر والواحد (٠ - ١). وتكون قيمة معامل التصحيح تساوي صفراً عندما يكون حجم الدراسة مساوياً لحجم العينة. ويحدث ذلك عندما ندرس جميع عناصر مجتمع الدراسة وفي هذه الحالة لا نستطيع حساب قيمة الخطأ المعياري. ويعتبر الخطأ المعياري ذو قيمة اذا تم دراسة العينات فقط وليس عند دراسة مجتمع الدراسة بالكامل. وتقترب قيمة معامل التصحيح من (١) عندما يكون حجم المجتمع (ن) كبيراً بالنسبة لحجم العينة (ن). لذلك ان عملية اضافة معامل التصحيح الى المعادلة سوف يقلل دائماً من قيمة الخطأ المعياري ما عدا الحالة التي يكون فيها قيمته (معامل التصحيح) مساوية للواحد. فاذا ما تمّ اهمال قيمة معامل التصحيح فان هذا يؤدي الى المبالغة في قيمة الخطأ المعياري وبالتالي زيادة طول فترة الثقة. وهذا بالتالي يؤدي الى التحيّز في ايجاد حدود فترة الثقة.

ـ تحديد حجم العينة The Determination of Sample Size

يجب الاهتمام الآن وبعد فهم أخطاء العينة وأخطاء غير العينة بسؤال تحديد حجم العينة.

- حجم العينة والنظرية الاحصائية.

لقد قمنا بحساب فترة الثقة للقيمة الاحصائية في حالة اختيار عينة عشوائية بسيطة معلومة الحجم. ولحساب فترة الثقة لمقياس فاننا بحاجة الى المعلومات التالية والتي تم تغطيتها خلال هذا الفصل.

١ - تقدير لقيمة الوسط، سَ .

٢ - تقدير لقيمة الانحراف المعياري، S .

٣ - حجم العينة.

٤ - مستوى الثقة.

٥ - استخدام (٢) و (٣) في حساب الخطا المعياري ($S\bar{X}$) . ومن ثم قمنا بحساب فترة الثقة المناسبة. والمعادلة المستخدمة لايجاد فترة الثقة عند مستوى ثقة مقداره ٩٥٪ كانت

$$\text{فترة الثقة} = \text{سَ} \pm ٢ \frac{S}{\sqrt{\text{ن}}}$$

اننا نستطيع ايجاد فترة الثقة اذا ما أوجدنا قيمة سَ و S بالاضافة الى معرفة حجم العينة. أو يمكن ان نجد قيمة الدقة في التقدير عن طريق استخدام جزء من معادلة فترة الثقة وكما يلي :

$$\text{الدقة} = \pm ٢ \frac{S}{\sqrt{\text{ن}}}$$

افرض الآن اننا نريد تحقيق مستوى معين من الدقة. فاذا كان معلوم لدينا قيمة الانحراف المعياري للمتغير (S) فاننا نستطيع حل المعادلة بالنسبة لحجم العينة المطلوب.

دعنا الآن نقوم بتوضيح هذه النتيجة. افرض االنا نريد الحصول على تقدير لمتوسط

عمر الطلبة بحيث تكون الدقة في التقدير ما بين ± ٣ سنوات عند مستوى ثقة قدره ٩٥٪ . وافرض اننا قمنا أيضاً بتقدير قيمة S وكانت S = ٣، فانه يمكن الحصول على حجم العينة المطلوب بعد حل المعادلة بالنسبة لحجم العينة (ن) .

$$\text{الدقة} = \pm ٢ \times \frac{S}{\sqrt{\text{ن}}}$$

$$\pm ٣ = \pm ٢ \times \frac{٣}{\sqrt{\text{ن}}}$$

$$٣ = \frac{٦}{\sqrt{\text{ن}}}$$

$$٣\sqrt{\text{ن}} = ٦$$

$$\sqrt{\text{ن}} =$$

$$\sqrt{\text{ن}} = ٢$$

$$\text{ن} = ٤$$

وهذا يعني انه اذا كان حجم العينة ن = ٤، فان هذا سوف يضمن دقة في التقدير قدرها ± ٣ سنوات اذا كانت قيمة S = ٣.

وقد قمنا بالتعبير عن الدقة في مثالنا بوحدة السنوات. وعليه، فانه يطلق على الدقة في هذه الحالة بالدقة المطلقة. كما يمكن التعبير عن الدقة يضاً بنسبة قيمة الوسط الذي تم حسابه. وفي هذه الحالة تتغير الدقة بناءاً على قيمة الوسط. وتدعى الدقة التي تمّ التعبير عنها بالنسب بالدقة النسبية. دعنا نقوم بحل مثال لحساب حجم العينة لضمان نسبة دقة محددة. افرض ان سَ = ٢٥ ، S = ٣ ، نسبة الدقة المطلوبة هي ± ١٠٪ (١ر)، ومستوى الثقة المرغوب به = ٩٩٫٧٪ فان حجم العينة المطلوب لتحقيق نسبة الدقة هو

$\pm$ ب سَ $= \pm 3 \frac{S}{\sqrt{ن}}$ ، حيث ب = نسبة الدقة المطلوبة

$\pm (1ر) (25) = \pm 3 \times \frac{3}{\sqrt{ن}}$

$2ر5 = \frac{9}{\sqrt{ن}}$

$2ر5 \sqrt{ن} = 9$

$\sqrt{ن} = \frac{9}{2ر5}$

$\sqrt{ن} = 3ر6$

$ن = 13$

وهذا يعني ان حجم العينة المطلوب هو ١٣ اذا كانت سَ = ٢٥ و $S = 3$.

ويمكن اعادة كتابة معادلة نسبة الدقة بعد اهمال الاشارة $\pm$ لانها تسقط في الحسابات وكما يلي :

$ب \times سَ = \frac{9}{\sqrt{ن}}$

$ب\ سَ \sqrt{ن} = 3 S$

$ب \sqrt{ن} = 3 \frac{S}{\sqrt{سَ}}$

ان اعادة كتابة المعادلة كما في الشكل اعلاه يبين الحاجة لمعرفة كل من سَ و S لحساب حجم العينة، انما معرفة نسبة الانحراف المعياري الى الوسط الحسابي S / سَ تكفي لايجاد حجم العينة. وغالباً ما يطلق على هذه النسبة بمعامل التغير.

مثال ١ : أوجد حجم العينة لمتغير غير متصل اذا علمت ان قيمة ρ = ٣ر ، والدقة المطلقة التي يرغب بها الباحث هي ± ٠٤و اذا علمت ان مستوى الثقة يساوي ٩٥٪ ؟

$$\text{الدقة المطلوبة} = \pm\ ٢ \sqrt{\frac{q\rho}{ن}}$$

$$٠٤ر = ٢ \sqrt{\frac{(٣ر)\ (٧ر)}{ن}}$$

$$٠٤ر = ٢ \sqrt{\frac{٢١ر}{ن}}$$ ، بالقسمة على ٢

$$٠٢ر = \sqrt{\frac{٢١ر}{ن}}$$

$$٠٢ر \sqrt{ن} = \sqrt{٢١ر}$$

$$\sqrt{ن} = \frac{\sqrt{٢١ر}}{٠٢ر}$$

$$\sqrt{ن} = \frac{٤٦ر}{٠٢ر}$$

$$\sqrt{ن} = ٢٣$$

$$ن = ٥٢٩$$

وهذا يعني ان حجم العينة المطلوب هو ٥٢٩ .

مثال ٢ : أوجد حجم العينة لمتغير غير متصل اذا علمت ان قيمة ρ = ٣، ونسبة الدقة المطلوبة كانت ± ٥٪ اذا علمت ان مستوى الثقة يساوي ٧ر٩٩٪ .

$$\text{نسبة الدقة} = ٣\sqrt{\frac{pq}{\text{ن}}}$$

$$\text{ب}\,p = ٣\sqrt{\frac{pq}{\text{ن}}}$$

$$(٠٫٠٥)(٠٫٣) = ٣\sqrt{\frac{(٠٫٣)(٠٫٧)}{\text{ن}}}$$

$$٠٫٠١٥ = ٣\sqrt{\frac{٠٫٢١}{\text{ن}}} \quad \text{، بالقسمة على ٣}$$

$$٠٫٠٠٥ = \sqrt{\frac{٠٫٢١}{\text{ن}}}$$

$$٠٫٠٠٥\sqrt{\text{ن}} = \sqrt{٠٫٢١}$$

$$\sqrt{\text{ن}} = \sqrt{٠٫٢١} \,/\, ٠٫٠٠٥$$

$$\sqrt{\text{ن}} = ٩٢$$

$$\text{ن} = ٨٤٦٤ \,.$$

وهذا يعني ان حجم العينة المطلوب هو ٨٤٦٤ .

لاحظ ان حجم العينة الذي تم حسابه لحد الآن كان لعينات عشوائية بسيطة فقط. فاذا كانت اجراءات اختيار العينة أكثر تعقيد من العشوائية البسيطة كالطبقية أو العنقودية، فان ذلك يؤدي الى استخدام معادلة كثر تعقيد لايجاد حجم العينة. ولكن مبدأ عمل جميع المعادلات هو نفس المبدأ. بمعنى آخر، حتى نتمكن من ايجاد حجم العينة لعينات احتمالية غير العشوائية البسيطة فانه لا بد من تحديد مستوى الثقة والدقة المطلوبة كما هو الحال بالنسبة للعينات العشوائية البسيطة ومن ثم نقوم باستخدام المعادلة المناسبة في قياس الخطأ المعياري لنتمكن من ايجاد حجم العينة.

- حجم العينة والخطأ غير العيني

لا يجوز لأي شخص ان يقبل حجم العينة المحسوب بالمعادلة الاحصائية بدون تفكير. ان احد الأسباب التي لا تنصح بعمل ذلك هو وجود الأخطاء غير العينية. ان بعض الأخطاء غير العينية تصبح ذات قيم أكبر كلما زاد حجم العينة مثل أخطاء عدم الاستجابة وأخطاء المقابلة وأخطاء اعداد البيانات بالاضافة الى أخطاء تحليل البيانات. لذلك فالدراسة التي تنفذ بدقة على عينة حجمها (٢٠٠) فرداً قد تحتوي على أخطاء اقل مقارنة مع دراسة تنفيذ بنفس الدقة على عينة حجمها ٢٠٠٠ والسبب في ذلك زيادة احتمال الوقوع في الأخطاء غير العينية. ان الاعتقاد السائد لدى عامة الناس بأنه كلما كان حجم العينة كبير كلما كان افضل من العينة الصغيرة والذي هو بالواقع غير صحيح. لذلك يتوجب على الباحثين مواجهة هذا التفكير بالنسبة لمجتمع الدراسة وان لا ينخدعوا بهذا التفكير او الاعتقاد لنفس الأسباب. فالكبر لا يعني بالضرورة الأفضل. ان بعض العينات الصغيرة تؤدي الى مستوى من الدقة الاحصائية بحيث يكون مقنع لاصحاب القرار.

- حجم العينة والعوامل الأخرى

يجب على أي دراسة بحث ان توازن بين الأناقة في خطوات البحث والمعضلات العملية التي يمكن ان تواجه الباحث. وهذه المعضلات قد تؤثر على قرار حجم العينة بالاضافة الى قرارات أخرى سوف نقوم بشرحها.

١ - أهداف الدراسة.

ان نوع الاستخدام للمعلومات التي يحصل عليها الشخص الباحث من الدراسة تؤثر على حجم العينة. فالقرارات التي لا تحتاج الى دقة بالمعلومات يمكن عملها بعينة ذات حجم صغير. فيمكن لشركة ان تكون سعيدة في قياس رغبة المستهلكين بسلعة جديدة من خلال نسبة قدرها ١٥٪ الى ٢٠٪ . اما اذا كانت

القرارات تحتاج الى دقة في المعلومات فانه لا يمكن عمل هذه القرارات من خلال دراسة عينة صغيرة. اذا أراد سياسي توقع نتائج انتخابات برلمانية مثلاً بدقة احصائية معينة فانه لا بد ان يختار عينة ذات حجم كبير.

٢ ـ مشكلة الوقت.

اذا كانت نتيجة البحث مطلوبة في وقت قصير فاننا لا نتوقع استخدام أي عينة سوى العينة ذات الحجم القليل. لأنه كلما كبر حجم الدراسة كلما زاد الوقت المتوقع ان يحتاجه الباحث لتنفيذ الدراسة.

٣ ـ مشكلة التكاليف.

ان مقدار الأموال المخصصة لعمل دراسة البحث ذات تاثير على حجم العينة التي يجب أن يختارها الباحث. وهذا لا يعني انه اذا كانت الأموال المخصصة لعمل بحث كثيرة فانه يجب عمل دراسة البحث على عينة كبيرة الحجم. ولكن يجب دائماً أن لا ننسى حجم العينة المطلوب لتحقيق أهداف دراسة البحث.

٤ ـ اجراءات تحليل البيانات.

ان اجراءات تحليل البيانات ذات تاثير على حجم العينة. ان معظم الأنواع الأساسية في التحليل تتعامل مع متغير واحد فقط في كل مرحلة تحليل. ويدعى هذا النوع من التحليل بالتحليل أحادي المتغير. ولكن في بعض الحالات قد نحتاج الى استخدام اختبارات احصائية ثنائية المتغيرات أو متعدد المتغيرات لتنفيذ الدراسة قيد البحث. يجب ملاحظة هنا ان كل اختبار من الاختبارات الاحصائية سواء كان أحادي المتغير أو ثنائي أو متعدد المتغيرات فانه يحتاج الى حجم عينة مختلف حتى يتمكن الشخص الباحث من الحصول على تقديرات صالحة لمعالم مجتمع الدراسة. وبشكل عام يتحدد حجم العينة بناءاً على عدد

معالم مجتمع الدراسة التي نحتاج الى تقديرها. فكلما زاد عدد معالم مجتمع الدراسة التي نحتاج الى تقدير كلما زادت الحاجة الى عينة ذات حجم كبير.

مما تقدم نلاحظ ان عملية تحديد حجم العينة الأمثل يعتمد على أكثر من متغير. من هذه المتغيرات حجم الدقة الاحصائية المطلوبة ودرجة الاهتمام بالأخطاء غير العينية وأهداف الدراسة والوقت المتاح والتكاليف هذا بالاضافة الى خطة تحليل البيانات. لذلك لا يوجد هناك اجابة واحدة صحيحة بالنسبة لاختيار حجم العينة الأمثل لأي دراسة.

أسئلة

س١ : ما هو الفرق اذا كان هناك ما بين المعلمة والقيمة الاحصائية ؟

س٢ : ما هي الرموز المستخدمة للتعبير عن المعالم والقيم الاحصائية لكلا النوعين من المتغيرات المتصلة والمتقطعة ؟

س٣ : عرف كل من الوسط والتباين بالكلمات ثم بالرموز الرياضية لكل من المجتمع والعينة مرة لمتغير متصل ومرة لمتغير متقطع.

س٤ : لماذا تعتبر نظرية الحد المركزية مهمة في قياس الخطأ العيني ؟

س٥ : ما هو المقصود بفترة الثقة ؟

س٦ : ما هو تأثير حجم العينة على فترة الثقة ؟

س٧ ما هي الخصائص المرغوب توفرها في القيم الاحصائية ؟

س٨ : ما هي المشاكل التي يمكن ان تواجه اي دراسة ميدانية تستخدم العينة العشوائية البسيطة ؟

س٩ : ما هو المقصود بالدقة المطلقة، الدقة النسبية ؟

س١٠ : ما هي المعلومات الضرورية الواجب توفرها لحساب الحجم الأمثل للعينة العشوائية ؟

س ١١ : أذكر العوامل التي يجب ان تؤخذ بعين الاعتبار لتحديد حجم العينة لدراسة ما ؟

الفصل الخامس

طرق العينات الأكثر تعقيداً

More Complex Sampling Procedures

طرق العينات الأكثر تعقيداً
More Complex Sampling Procedures

(١) العينة الطبقية

ان أحد الخصائص التي نرغب وجودها في التقديرات الاحصائية هي الكفاءة. بمعنى آخر اننا تريد أن يكون الخطأ المعياري في التقديرات الاحصائية أقل ما يمكن. ان استخدام العينة الطبقية والعنقودية ممكن أن يؤدي الى التقليل من الخطأ المعياري للتقدير الاحصائي. وعليه فاننا نتوقع أن تكون عرض فترة الثقة أقل. ولكن تعتبر اجراءات اختيار العينة سواء كانت الطبقية أو العنقودية أكثر تعقيداً من اجراءات اختيار العينة العشوائية البسيطة. ومع ذلك فان هاتين الطريقتين مستخدمتين بالحياة العملية بشكل واسع.

ـ طريقة اختيار العينة الطبقية

ان عملية اختيار العينة الطبقية تتم كما يلي :

١ ـ يجب تقسيم مجتمع الدراسة الى طبقات أو مجموعات مانعة لبعضهما البعض. والمقصود بمانعة لبعضهما البعض هو عدم وجود اي عنصر في أكثر من طبقة. مثال: يمكن أن يتم تقسيم مجتمع الدراسة الى طبقات حسب الجنس. وعليه فاننا نتوقع الحصول بعد تقسيم مجتمع الدراسة على طبقتين : طبقة للذكور وطبقة للاناث. كما اننا لا نتوقع انتماء أي فرد من افراد المجتمع الى كلا الطبقتين. كما يجب أن تحتوي الطبقات المختلفة على جميع عناصر مجتمع الدراسة. بمعنى آخر يجب على طبقة الذكور وطبقة الاناث تعريف جميع عناصر مجال متغير الجنس

بحيث لا يمكن الحصول على اي طبقة اخرى بالنسبة لمتغير الجنس.

٢ - اختيار عينة عشوائية مستقلة من كل طبقة.

- توضيح عمل طريقة العينة الطبقية.

تعتبر هذه الطريقة عملية جداً في تقليل الخطأ المعياري للتقدير الاحصائي اذا تم تقسيم المجتمع الى طبقات بحيث يكون التجانس عالي في الطبقة الواحدة بالنسبة للمتغير ذو العلاقة بالحسابات الاحصائية. اما اذا كان التجانسي في المجتمع الاحصائي ككل افضل منه في كل طبقة على حده فاننا سوف لا نتوقع اي تقليل من وقوع الخطأ المعياري.

وسوف نستخدم جزء من البيانات المستخدمة في العينة العشوائية والبسيطة والمعروضة في الجدول (٤ - ٢) والتي تتعلق بالخانة الأولى من الرقم الجامعي والعمر لتوضيح خاصية العينة الطبقية في تقليل الخطأ المعياري. وجدول (٥ - ١) يعرض الخانة الأولى من الرقم الجامعي والعمر للطلبة. وسوف نستخدم متغير العمر لتوضيح خاصية تقليل الخطأ المعياري.

جدول (٥ ـ ١)

	١ ـ الخانة الأولى من الرقم الجامعي	٢ـ العمر (س ١)	١ ـ الخانة الأولى من الرقم الجامعي	٢ـ العمر (س ١)	
	١-	٢٥	٢٦-	٢٢	
	٢-	٢٧	٢٧-	١٩	
	٣-	٢٩	٢٨-	٢٠	
	٤-	٣١	٢٩-	١٩	
	٥-	٢٥	٣٠-	٢٤	
	٦-	٢٩	٣١-	٢٥	
	٧-	٢٧	٣٢-	٢٢	
	٨-	٢٤	٣٣-	٢٠	
	٩-	٢٧	٣٤-	٢١	
	١٠-	٢٨	٣٥-	٢١	
	١١-	٢٣	٣٦-	٢٣	
	١٢-	٢٩	٣٧-	٢١	
	١٣-	٢٦	٣٨-	٢٣	
طلبة الدراسات العليا	١٤-	٢٨	٣٩-	١٨	طلبة البكالوريوس
	١٥-	٢٨	٤٠-	٢١	
	١٦-	٢٦	٤١-	١٩	
	١٧-	٢٦	٤٢-	٢٣	
	١٨-	٣٦	٤٣-	٢٢	
	١٩-	٢٨	٤٤-	١٩	
	٢٠-	٢٦	٤٥-	٢٠	
	٢١-	٢١	٤٦-	٢٠	
	٢٢-	١٩	٤٧-	٢١	
	٢٣-	٢٤	٤٨-	٢٠	
	٢٤-	٢٢	٤٩-	١٩	
	٢٥-	٢٠	٥٠-	١٨	

نلاحظ من الجدول ان التجانس ما بين أعمار طلبة الدراسات العليا (١ - ٢٠) عالي، وكذلك نفس الشيء بالنسبة لطلبة دراسة البكالوريوس من ٢١ — ٥٠ . وباستخدام المتغير الطبقي (دراسات عليا ضد دراسات بكالوريوس) فقد حصلنا على طبقتين اكثر تجانساً مقارنة مع المجتمع الذي تم الحصول عليه بناءاً على متغير العمر. لذلك فقد قمنا بعرض السبب الضروري والمهم للاستفادة من تقسيم المجتمع الى طبقات.

افرض اننا قمنا بسحب عينة عشوائية بسيطة بعنصرين من طبقة طلبة الدراسات العليا وعينة عشوائية بسيطة بثلاثة عناصر من طبقة طلبة الدراسات الدنيا. ان نسبة العينة الأولى الى الثانية هي ٢ : ٣ وهذا متناسق مع نسبة مجتمع طلبة الدراسات العليا الى طلبة دراسات بكالوريوس ٢٠ / ٣٠. ويطلق على هذا النوع من العينات بالعينات الطبقية المتناسقة. ويحدث هذا النوع من العينات عندما يتم اختيار عناصر العينة من كل طبقة على أساس نسبة عناصر الطبقة الواحدة الى عناصر مجتمع الدراسة. وللباحث الحق في اختيار عناصر العينة بشكل لا يتناسق مع توزيع المجتمع في كل طبقة.

دعنا الآن نقوم بتحديد بعض الرموز لتسهيل الشرح اللاحق :

$ن_{ط١}$ حجم المجتمع في الطبقة الأولى. $\bar{س}_{ط١}$ وسط العينة في الطبقة الأولى.

$ن_{ط٢}$ حجم المجتمع في الطبقة الثانية. $\bar{س}_{ط٢}$ وسط العينة في الطبقة الثانية.

$ن_{ط١}$ حجم العينة في الطبقة الأولى. $S^{٢}_{ط١}$ تباين العينة في الطبقة الأولى.

$ن_{ط٢}$ حجم العينة في الطبقة الثانية. $S^{٢}_{ط٢}$ تباين العينة في الطبقة الثانية.

افرض اننا قمنا الآن بسحب عينة عشوائية بسيطة حجمها عنصرين من الطبقة الأولى وعينة حجمها (٣) عناصر من الطبقة الثانية. وعليه فان شكل العينية الطبقية سوف يكون كما يلي :

الخانة الأولى من الرقم الجامعي	العمر
الطبقة الأولى	
٥	٢٥
١٧	٢٦
الطبقة الثانية	
٣٢	٢٢
٣٧	٢١
٤١	١٩

لغايات المقارنة مع نتائج العينة العشوائية البسيطة، فقد قمنا باختيار نفس عناصر العينة العشوائية البسيطة.

ـ حساب القيم الاحصائية لكل طبقة.

لقد قمنا في الفصل الرابع بحساب كل من الوسط والتباين والانحراف المعياري لجميع عناصر العينة وبدون تقسيم. وكانت القيم الاحصائية كما يلي

$\bar{س}$ = ٢٢٫٦ ، S^2 = ٨٫٣ ، S = ٢٫٨٨

دعنا الآن نقوم بحساب الوسط والتباين والانحراف المعياري لكل طبقة.

بالنسبة للطبقة الأولى :

$$\text{الوسط} = \bar{س}_{ط١} = \frac{\sum_{ر=١}^{ن_{ط١}} س_ر}{ن_{ط١}}$$

$$= \frac{٢٥ + ٢٦}{٢}$$

$$= \frac{٥١}{٢}$$

$$= ٢٥ر٥$$

ولحساب التباين $S^{٢}_{ط١}$

$$S^2 = \frac{\sum_{ر=١}^{ن_{ط١}} س_ر^{٢} - (\sum_{ر=١}^{ن_{ط١}} س_ر)^{٢} / ن}{ن - ١}$$

$$\sum س^{٢}_{ط١} = (٢٥)^{٢} + (٢٦)^{٢} = ٦٢٥ + ٦٧٦ = ١٣٠١$$

$$\frac{(\sum س_{ط١})^{٢}}{ن} = \frac{(٥١)^{٢}}{٢} = \frac{٢٦٠١}{٢} = ١٣٠٠ر٥$$

$$S^{٢}_{ط١} = \frac{١٣٠١ - ١٣٠٠ر٥}{١}$$

$$S^{٢}_{ط١} = ر٥$$

وعليه يكون الانحراف المعياري $S_{ط١} = ٠ر٧١$

أما بالنسبة للعينة الثانية

الوسط = سَ ط٢ = ٢٢ + ٢١ + ١٩ / ٣

$$سَ_{ط٢} = \frac{٦٢}{٣}$$

$$سَ_{ط٢} = ٢٠ر٧$$

ولحساب التباين نجد ما يلي

$$\sum س^{٢}_{ط٢} = (٢٢)^{٢} + (٢١)^{٢} + (١٩)^{٢}$$

$$= ٤٨٤ + ٤٤١ + ٣٦١$$

$$= ١٢٨٦$$

$$\frac{(\sum س_{ط٢})^{٢}}{ن} = \frac{(٦٢)^{٢}}{٣}$$

$$= \frac{٣٨٤٤}{٣}$$

$$= ١٢٨١ر٣$$

$$S^{٢}_{ط٢} = \frac{١٢٨٦ - ١٢٨١ر٣}{٢}$$

$$= \frac{٤ر٧}{٢}$$

$$= ٢ر٣٥$$

وعليه تكون قيمة S = ١ر٥٣ .

والجدول (٥ - ٢) يلخص هذه النتائج

نلاحظ من الحسابات ان التباين والانحراف المعيراي لكل طبقة أقل بكثير من التباين والانحراف المعياري للعينة ككل. وهذا يظهر واحدة من ايجابيات العينة الطبقية. وسوف ينتج عن ذلك ان طول فترة الثقة لكل طبقة سوف تكون أقصر من طول فقرة الثقة بالنسبة لجميع العناصر ككل .

جدول (٥ - ٢)

الوسط والتباين والانحراف المعياري في كل طبقة وللعينة ككل

	الوسط	التباين	الانحراف المعياري
بدون طبقات	٢٢ر٦	٨ر٣	٢ر٨٨
الطبقة الأولى	٢٥ر٥	ر٥	ر٧١
الطبقة الثانية	٢٠ر٧	٢ر٣٥	١ر٥٣

ـ حساب الوسط والخطأ المعياري للعينة ككل.

يجب أن يتحول الاهتمام الآن الى كيفية حساب الوسط والخطأ المعياري للعينة ككل بناءاً على نتائج الطبقات المختلفة. فوسط العينة سَ هو ببساطة عبارة عن معدل أوساط الطبقات المختلفة. وان وزن الطبقة هو نسبة حجم الطبقة الى حجم المجتمع الكلي. ويمكن التعبير عنه رياضياً كما يلي $ن_{ط ر} / ن$. حيث ان ن ط ر تمثل عدد الأشخاص في الطبقة ر. لذلك

$$سَ_{للطبقات} = \sum_{ر=١}^{أ} \left(\frac{ن_{ط ر}}{ن} \right) سَ_{ط ر}$$

حيث ان أ تمثل عدد الطبقات

ففي مثالنا الحالي فقد تم تقسيم المجتمع الى طبقتين، لهذا فان أ = ٢ ، وعليه

$$\bar{س}(كلي) = \sum_{ر=1}^{2} \left(\frac{ن_{طر}}{ن}\right) \bar{س}_{طر} = \left(\frac{ن_{ط1}}{ن}\right) \bar{س}_{ط1} + \left(\frac{ن_{ط2}}{ن}\right) \bar{س}_{ط2}$$

$$= \left(\frac{٢٠}{٥٠}\right) \times (٢٥٫٥) + \left(\frac{٣٠}{٥٠}\right) \times (٢٠٫٧)$$

$$= (٫٤)(٢٥٫٥) + (٫٦)(٢٠٫٧)$$

$$= ١٠٫٢ + ١٢٫٤$$

$$= ٢٢٫٦$$

نلاحظ ان الوسط المحسوب عن طريق استخدام العينة الطبقية هو نفس الوسط الذي حصلنا عليه دون استخدام العينة الطبقية.

ان نسبة $ن_{طر} / ن$ هي وزن النسبة المرافق لكل طبقة. ودعنا نرمز لها من الآن بـ $و_ر$.

ففي مثالنا يمكن القول بأن

$و_1 = ٫٤ ، و_2 = ٫٦$

ان حساب الخطأ المعياري للوسط باستخدام الطبقية أكثر تعقيداً مقارنة مع استخدام العشوائية البسيطة. والخطأ المعياري للوسط يساوي الجذر التربيعي لأوزان مربع الخطأ المعياري في كل طبقة. وعامل الوزن في هذه الحالة يساوي مربع الوزن النسبي لكل طبقة. وعليه فان المعادلة

$$S_{\bar{س}} = \sqrt{S^2_{\bar{س}}} \quad \text{(تمّ اهمال المجتمعات المحدودة)}$$

$$S^{2}_{\bar{س}} = \sum_{ر=1}^{أ} \left(\frac{ن_{طر}}{ن}\right) S^{2}_{\bar{س}_{طر}}$$

$$= \sum_{ر=1}^{أ} و_{ر}^{2}\, S^{2}_{طر}$$

ففي مثالنا الحالي، فان الخطأ المعياري بالنسبة للطبقة الأولى هو

$$S_{\bar{س}_{ط1}} = \frac{ر٧١}{\sqrt{٢}} = \frac{ر٧١}{١ر٤١}$$

$$= ر٥$$

اما $S^{2}_{\bar{س}_{ط1}} = ر٢٥$

اما بالنسبة للطبقة الثانية

$$S_{\bar{س}_{ط2}} = \frac{١ر٥٣}{\sqrt{٣}} = \frac{١ر٥٣}{١ر٧٣}$$

$$= ر٨٨$$

اما $S^{2}_{ط2} = ر٧٨$

لهذا

$$S^{2}_{\bar{س}} = (ر٤)^{٢}(ر٢٥) + (ر٦)^{٢}(ر٧٨)$$

$$= (ر١٦)(ر٢٥) + (ر٣٦)(ر٧٨)$$

$$= ر٠٤ + ر٢٨$$

$$= ر٣٢$$

اما

$$S_{\bar{\text{س}}} = \sqrt{\text{٣٢ر}}$$

$$= \text{٥٧ر٠}$$

لاحظ اننا قد قمنا بتربيع معامل الوزن لكل طبقة ومربع الخطأ المعياري.

ويمكن احتساب $S^2_{\bar{\text{س}}}$ مباشرة ودون حساب الخطأ المعياري لكل طبقة. والمعادلة التي يمكن أن تستخدم لعمل ذلك هي :

$$S^2_{\bar{\text{س}}} = \sum_{\text{ر}=1}^{\text{أ}} \frac{\left(\frac{\text{ن}_{\text{طر}}}{\text{ن}}\right)^2 S^2_{\bar{\text{س}}_{\text{طر}}}}{\text{ن}_{\text{طر}}}$$

$$= \sum_{\text{ر}=1}^{\text{أ}} \frac{(\text{و}_{\text{ر}})^2 S^2_{\text{طر}}}{\text{ن}_{\text{طر}}}$$

ففي مثالنا الحالي ، أ = ٢

$$S^2_{\bar{\text{س}}} = \frac{(\text{و}_1)^2 S^2_{\text{ط}1}}{\text{ن}_{\text{ط}1}} + \frac{(\text{و}_2)^2 S^2_{\text{ط}2}}{\text{ن}_{\text{ط}2}}$$

$$= \frac{(\text{٤ر})^2 (\text{٥ر})}{\text{٢}} + \frac{(\text{٦ر})^2 (\text{٢ر٣٥})}{\text{٣}}$$

$$= \frac{\text{١٦ر} \times \text{٥ر}}{\text{٢}} + \frac{\text{٣٦ر} \times \text{٢ر٣٥}}{\text{٣}}$$

$$= \text{٠٤ر} + \text{٢٨ر}$$

$$= \text{٣٢ر}$$

أما بالنسبة $S_{\bar{س}} = \sqrt{\text{٣٢ر}}$

= ٥٧ر

نلاحظ ان قيمة $S_{\bar{س}}$ هي نفس القيمة المحسوبة سابقاً. ان قيمة الخطأ المعياري للوسط التي تم حسابها بدون التقسيم الى طبقات كانت (٣ر١). ولكن تناقصت قيمة الخطأ المعياري للوسط بعد التقسيم الى طبقات من (٣ر١) الى (٥٧ر) .

ـ حساب فترات الثقة المصاحبة.

ان فترة الثقة الـ ٩٥٪ للعينة الطبقية هي :

٢٢ر٦ ± ٢ (٥٧ر) = ٢٢ر٦ ± ١٠١

= ٢١ر٥ – ٢٣ر٧

بينما كانت فترة الثقة الـ ٩٥٪ للعينة غير الطبقية هي

٢٠ – ٢٥ر٢

لذلك نلاحظ ان طول فترة الثقة قد تناقص من ٢ر٥ الى ٢ر٢ وكذلك تناقصت الدقة المطلقة من ± ٢ر٦ الى ± ١ر١ . لاحظ الآن ان وسط المجتمع μ = ٢٣ر٧ قد وقع بين حدود فترة الثقة الـ ٩٥٪ الجديدة. كما نلاحظ ان فترة الثقة الـ ٩٩ر٧٪ للعينة الطبقية قد تناقصت مقارنة مع فترة الثقة للعينة غير الطبقية. ان فترة الثقة الـ ٩٩ر٧٪ للعينة غير الطبقية هي

٢٢ر٦ ± ٣ (٥٧ر) = ٢٢ر٦ ± ١ر٧ =

٢٠ر٩ – ٢٤ر٣

نلاحظ مما تقدم ان اجراءات العينة الطبقية أكثر فعالية من العينة غير الطبقية. ان هذا يبرر الاستخدام الواسع للعينة الطبقية في الحياة العملية.

ان السبب في الحصول على تناقص في الخطأ المعياري والدقة المصاحبة في استخدام العينة الطبقية هو استخدام التغير في الطبقة الواحدة في حساب الخطأ المعياري الكلي. وبذلك يصبح التنغير ما بين الطبقات تغيراً غير ضروري في حساب الخطأ المعياري الكلي.

١ ـ عدد العينات الممكنة

افرض ان هناك مجتمع دراسة مكون من خمسة عناصر هي أ، ب، جـ د، هـ. وافرض ان العنصرين أ و ب من طبقة والعناصر جـ، د، هـ من طبقة أخرى. اذا كان الباحث يرغب في اختيار عينة مكونة من عنصرين بحيث يكون كل عنصر من طبقة. ان العنصر الذي يمكن اختياره من الطبقة الأولى هو اما أ أو ب. بينما العنصر الذي يمكن اختياره من الطبقة الثانية هو اما جـ أو د أو هـ. ان تجميع العناصر الممكن اختيارها من كل طبقة يؤدي الى تشكيل العينات الممكنة التالية.

رقم العينة	عناصر العينة
١	أ جـ
٢	أ د
٣	أ هـ
٤	ب جـ
٥	ب د
٦	ب هـ

نلاحظ ان عدد العينات الممكن تشكيلها باستخدام الطبقية هو (٦) عينات مختلفة بينما عدد العينات الممكن بدون الطبقية هو (١٠). ان العينات الممكنة التي لم تظهر هي تلك العينات التي يمكن ان تقع بنفس الطبقة. فعلى سبيل المثال :

فالعينتان أ ب و جـ د لم تعد عينات ممكنة لأن العنصرين أ و ب يقعان في الطبقة الأولى والعنصرين جـ و د يقعان في الطبقة الثانية.

نلاحظ مما سبق ان عدد العينات الممكنة عبارة عن حاصل ضرب عدد جميع العينات الممكنة في كل طبقة.

ففي مثالنا الحالي، نجد ان عدد العينات الممكنة هو ٢ X ٣ = ٦.

كما يجب التنويه هنا الى ان التوزيع العيني لأوساط العينات الممكن اختيارها باستخدام الطريقة الطبقية سوف يشكل منحنى طبيعي. كما ان الخطأ المعياري باستخدام الطريقة الطبقية سوف يكون أقل من الخطأ المعياري باستخدام العينة العشوائية البسيطة.

٢ ـ العينة العنقودية.

ان اختيار عناصر العينة باستخدام طرق العينة الاحتمالية التي تم شرحها كان يتم على أساس فردي. إلا ان طريقة الاختيار في الطريقة العنقودية يختلف عن طريقة الاختيار سواء كان في الطريقة العشوائية البسيطة أو الطبقية، حيث يتم الاختيار بالطريقة العنقودية لمجموعة أو عنقود من العناصر بطريقة احتمالية مرة واحدة. لذلك يجب تقسيم المجتمع الى عناقيد او مجموعات مختلفة مانعة لبعضها البعض وشاملة لجميع عناصر المجتمع. وبعد ذلك نقوم باختيار عينة عشوائية من هذه المجموعات أو العناقيد.

افرض ان هناك مجتمع مكون من (٢٠) عنصراً ومقسم الى (٤) مجموعات متساوية من حيث عدد العناصر وكما يلي :

المجموعة	رقم عنصر المجتمع
١	١ ، ٢ ، ٣ ، ٤ ، ٥
٢	٦ ، ٧ ، ٨ ، ٩ ، ١٠
٣	١١ ، ١٢ ، ١٣ ، ١٤ ، ١٥
٤	١٦ ، ١٧ ، ١٨ ، ١٩ ، ٢٠

اذا أردنا اختيار عينة عشوائية مكونة من (١٠) أشخاص، فانه يمكن اختيار العناصر بصورة فردية باستخدام العينة العشوائية البسيطة. أو من الممكن اختيار مجموعتين من المجموعات الأربعة ومن ثم نستخدم جميع العناصر الموجودة في هاتين المجموعتين. ان الوضع الذي نختار فيه المجموعتين ومن ثم نستخدم جميع العناصر المتوفرة في هذه المجموعات يطلق عليه اسم العينة العنقودية ذو المرحلة الواحدة. أما اذا قمنا باختيار عينة عشوائية من العناصر الموجودة في المجموعات التي تمّ اختيارها، فاننا نطلق على هذا الوضع بالعينة العنقودية ذو المرحلتين. وفي كلا الحالتين (العينة العنقودية والعينة العشوائية البسيطة) تكون نسبة العينة الى المجتمع نفسها والتي تساوي (٥ر) .

ان الطريقة العنقودية تقوم على أساس اختيار مجموعات كاملة بعد تقسيم المجتمع الى مجموعات عشوائياً وليس على أساس اختيار عناصر مستقلة من كل مجموعة كما هو الحال في الطريقة الطبقية. فاذا تمّ تقسيم المجتمع الى مجموعات غير متجانسة مشابهة بالضبط لعدم تجانس المجتمع، فان أي مجموعة يتم اختيارها عشوائياً سوف تكون ممثلة للمجتمع. ولكن يعتبر هذا بالحياة العملية مثالياً ويصعب تحقيقه. لذلك يجب ملاحظة هنا ان المعيار المستخدم في تقسيم المجتمع بالطريقة العنقودية هو عكس المعيار المستخدم في الطريقة الطبقية تماماً. ففي الطريقة الطبقية فاننا نبحث عن مجموعات متجانسة بينما نبحث في الطريقة العنقودية عن مجموعات غير متجانسة.

ولكن السؤال الذي يطرح نفسه الآن، ما هو حجم الخطأ المعياري الناتج عن العينة العنقودية مقارنة مع حجم الخطأ المعياري الناتج عن العينة العشوائية البسيطة؟ ان الاجابة عن هذا السؤال تعتمد على درجة التشابه في عدم التجانس بالنسبة للعينات التي تم الحصول عليها مقارنة مع المجتمع وكما يلي :

١ ـ اذا كانت جميع المجموعات متشابهة في عدم التجانس وكما هو الحال في المجتمع، فان كلا الطريقتين يكون لهما نفس الخطأ المعياري.

٢ ـ اذا كان عدم التجانس في المجموعات أقل من عدم التجانس في المجتمع، فان الخطأ المعياري الناتج عن العينة العنقودية سوف يكون أكبر من الخطأ المعياري الناتج عن العينة العشوائية البسيطة.

ان السبب في الاشارة الى مقارنة الخطأ المعياري الناتج عن اجراءات العينة المختلفة هو لتقييم الفعالية الاحصائية لهذه الاجراءات المختلفة.

ففي الحياة العملية، غالباً ما يكون عدم التجانس في العينات العنقودية أقل من عدم التجانس في المجتمع. وهذا يعني ان الكفاءة الاحصائية للعينات العنقودية غالباً ما تكون أقل من الكفاءة الاحصائية للعينات العشوائية البسيطة. ولكن لأسباب تتعلق بالتكاليف نجد ان استخدام العينات العنقودية بالحياة العملية واسع جداً لان اجراءات اختيار العينات العنقودية غالباً ما تكون أرخص من اجراءات اختيار العينة بالطرق الأخرى بالنسبة لحجم العينة.

٣ ـ العينات المنتظمة Systematic Sampling

يقوم الباحث بالعينة المنتظمة باختيار كل م عنصر واقع ضمن اطار الاختيار بعد تحديد نقطة بداية عشوائية في أي مكان واقع ضمن مجال اختيار العنصر الأول. افرض اننا نريد اختيار عينة منتظمة مكون من (٥) طلاب من مجتمع الطلاب الممثل في الجدول (٥ ـ ١). نجد في هذا المثال ان م = ٥٠ / ٥ = ١٠، وبشكل عام فان م = نَ / ن.

ولاختيار العينة المنتظمة نقوم بما يلي :

١ ـ اختيار رقم عشوائي ما بين (٠) و (١٠). وسوف يكون هذا العدد هو نقطة البداية وأول عنصر في العينة.

٢ ـ نضيف (١٠) الى الرقم العشوائي فيكون هذا الرقم هو العنصر الثاني في العينة. ثم نضيف (١٠) للحصول على العنصر الثالث وهكذا.

فاذا كانت البداية العشوائية عند الرقم (٢)، فان العناصر التي تتضمنها العينة هي :

٢ ، ١٢ ، ٢٢ ، ٣٢ ، ٤٢ .

ان اختيار العناصر التي يجب ان تتضمنها العينة يكون آلياً بعد تحديد فترة العينة ونقطة البداية العشوائية. وتشكل هذه العناصر بعد اختيارها عنقوداً. وفي حالة مثالنا الذي يتكون من مجتمع عدد عناصره يساوي (٥٠)، فانه يمكن اختيار (١٠) عينات منتظمة ممكنة من حجم (٥) أشخاص. والسبب في ذلك ان كل عنقود يتم اختياره يحتوي على (١ / ١٠) من عناصر المجتمع.

وبما اننا نستخدم جميع عناصر العينة التي تم اختيارها بالعينة المنتظمة، فانه يطلق عليها اسم اجراء العينة العنقودية ذو المرحلة. ان الوسط للعينة المنتظمة يساوي مجموع عناصر العينة مقسوماً على عدد عناصر العينة. وبعد الرجوع الى الجدول (٥ ـ ١) نجد ان اعمار الطلبة الذين تم اختيارهم هي : ٢٧، ٢٩، ١٩، ٢٢، ٢٣. وعليه فان وسط العمر للعينة هو

$$\bar{س} = \frac{٢٧ + ٢٩ + ١٩ + ٢٢ + ٢٣}{٥}$$

$$\bar{س} = \frac{١٢٠}{٥}$$

$$\bar{س} = ٢٤$$

ولايجاد التباين، نقوم بما يلي :

$\sum$ س$_{ر}^{٢}$ = (٢٧)^{٢} + (٢٩)^{٢} + (١٩)^{٢} + (٢٢)^{٢} + (٢٣)^{٢}

= ٧٢٩ + ١٨١ + ٣٦١ + ٤٨٤ + ٥٢٩

= ٢٩٤٤

($\sum$ س$_{ر}$)^{٢}/ن = $\dfrac{(١٢٠)^{٢}}{٥}$

= $\dfrac{١٤٤٠٠}{٥}$

= ٢٨٨٠

وعليه

$S^{٢}$ = $\dfrac{٢٩٤٤ - ٢٨٨٠}{٥ - ١}$

= $\dfrac{٦٤}{٤}$

= ١٦

∴ الانحراف المعياري S = ٤

والخطأ المعياري هو $\dfrac{٤}{\sqrt{٥}}$ = $\dfrac{٤}{٢٫٢٤}$

= ١٫٧٩

وفترة الثقة الـ ٩٥٪ هي

٢٤ ± ٢ (١٫٧٩) = ٢٤ ± ٣٫٥٨

= ٢٠٫٤٢ – ٢٧٫٥٨

نلاحظ هنا ان الوسط الحقيقي وقع أيضاً ضمن حدود فترة الثقة الـ ٩٥٪ .

ان استخدام العينة المنتظمة غالباً ما يكون أوسع من العينة العشوائية البسيطة لسهولتها ولرخص تنفيذها. ان استخدام العينة المنتظمة يعفي الشخص الباحث من القفز الى الأمام أو الى الخلف ضمن اطار العينة حسب ما يقودنا اليه الرقم العشوائي المختار. كما ان استخدام العينة المنتظمة يعفي الباحث من التفكير بالعناصر المكررة. ولكن نجد ان هذه المشاكل تتحقق في العينة العشوائية البسيطة وبشكل مستمر. وبما ان العينة المنتظمة تعتبر أحياناً بديلاً عن العينة العشوائية البسيطة، فانه يمكن استخدامها لاختيار عناصر الطبقة في العينة المنتظمة.

ان الفائدة الأخرى التي تتميز بها العينة المنتظمة مقارنة مع العينة العشوائية البسيطة هي عدم ضرورة وجود الاطار العيني الكامل لاختيار العينة المنتظمة. فيمكن للشخص الباحث ان يختار كل خامس منزل مثلاً وبدن وجود قائمة تحتوي على كامل المنازل المتاحة.

أسئلة

س١ : كيف يتم اختيار العينة الطبقية ؟

س٢ : ما هو الهدف من وراء استخدام العينة الطبقية ؟

س٣ : ما هو المقصود بالعينات الطبقية المتناسقة ؟

س٤ : كيف يتم حساب الوسط والانحراف المعياري للعينة ككل من العينة الطبقية ؟

س٥ : أذكر الظروف التي تساعد العينة الطبقية على تقليل الخطا المعياري ؟

س٦ : لماذا تقلّل العينة الطبقية من الخطأ المعياري ؟

س٧ : ما هي العينة العنقودية ؟

س٨ : كيف يتم اختيار العينة المنتظمة ؟

الفصل السادس

عمليـة القيـاس

Measurement Process

عملية القياس
Measurement Process

- المقدمة

ان عملية القياس عملية أساسية وجانب رئيسي في البحث الاداري. وقد تكون عملية القياس أفضل طريقة لفهم الشيء أو الحدث الذي نريد قياسه. ولهذا السبب فان الاهتمام بموضوع القياس قد تزايد لجميع الباحثين في مجال الادارة.

ان لدى متخذي القرارات اهتمام كبير لقياس جوانب عديدة في النظام الاداري مثل الرضا الوظيفي، انتاج سلعة جديدة، تقسيم مجموعة المشترين حسب الخصائص الديمغرافية، اتجاهات العمال، اتجاهات المستهلك، تحديد فعالية استخدام نظام انتاجي جديد. هذا بالاضافة الى ان قياس الظاهرة الادارية أمر ضروري ومهم لعملية تزويد معلومات ذات معنى أو قيمة لمتخذي القرارات.

فشكل الأخطاء الناتجة عن عملية القياس جزء رئيسي بالنسبة لمعلومات الأبحاث الادارية. وفي معظم مشاريع الأبحاث تكون أخطاء الناتجة عن القياس أكبر من الأخطاء الناتجة عن اجراءات اختيار العينة وبشكل واضح. لذلك ان فهم مشاكل القياس وكيفية مراقبة هذه المشاكل يعتبر جانب مهم عند تصميم مشاريع الأبحاث الادارية بشكل فعّال.

وحتى يتمكن أي مدير من مراقبة أخطاء عملية القياس بشكل فعال يجب عليه الاهتمام بما يلي :

١ - ان تخصيص المعلومات المطلوبة قد يؤدي الى درجة من الصعوبة في تحديد مقياس دقيق.

٢ ـ وجوب ادراك اجراءات القياس البديلة في جمع المعلومات.

٣ ـ وجوب القيام بعملية تقييم تكلفة المقياس ضد دقة المقياس.

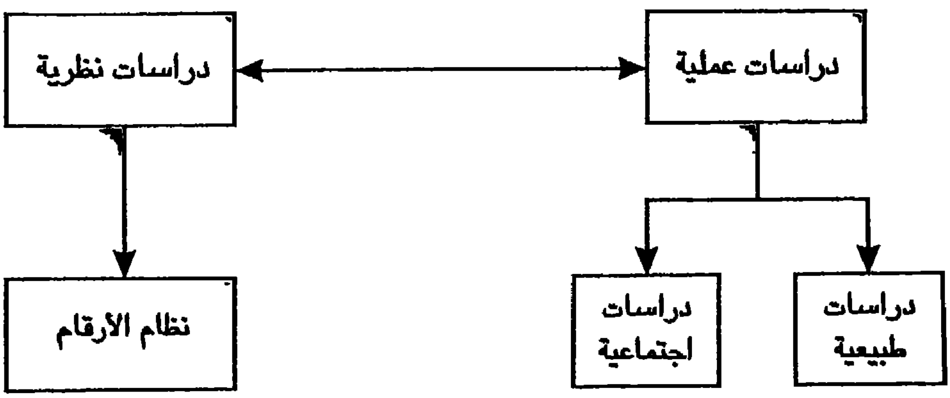
لذلك سوف يقوم هذا الفصل بالاهتمام بهذه الأمور.

ـ عملية القياس The Measurement Process

تتضمن عملية القياس في الأبحاث الادارية استخدام الأرقام لتمثيل الظاهرة الادارية قيد الدراسة. بمعنى آخر تحويل النظام العملي والذي يتمثل في دراسة الظاهرة الادارية الى نظام نظري يحتوي على أرقام فقط لتمثيل الظاهرة الادارية. والشكل (٦ ـ ١) يوضح عملية القياس كطريقة اتصال ما بين النظام العملي والنظام النظري. ويتكون النظام العملي من العلوم الطبيعية والعلوم الاجتماعية ذات العلاقة بالدراسات الانسانية. ان دراسات العلوم الادارية هي جزء من العلوم الاجتماعية لأنها تتضمن نشاطات انسانية موجهة لاشباع حاجات لا تتحقق الا عن طريق تبادل في العمليات.

شكل (٦ ـ ١)

عملية القياس



- تعريف القياس Definition of Measurement

ان الشرح السابق يقترح ان اهتمام عملية القياس هو في تطوير وسيلة اتصال ما بين النظام العملي (المتغير الاداري) والنظام النظري (الأرقام). لهذا يمكن تعريف القياس على انه عملية اعطاء أرقام الى خصائص الحاجات أو الأحداث باستخدام قوانين معينة. والمقياس الفعال يكون ممكن عندما تستجيب العلاقات الموجودة ما بين الحوادث أو المتغيرات في النظام النظري بشكل مباشر الى قوانين النظام الرقمي أو العددي. فاذا تحقق خطأ في تمثيل الاستجابة فان هذا يؤدي الى تحقق خطأ في القياس.

انه من المهم أن نلاحظ هنا ان تعريف القياس يقترح قياس خصائص الأحداث أو الحاجات أو المتغيرات وليس الأحداث أو المتغيرات نفسها. بمعنى آخر اننا لا نسعى لقياس آلة من الآلات المستخدمة في خط الانتاج أو سلعة أو عامل، ولكن نسعى لقياس خصائص الآلة أو السلعة أو العامل. لذلك فاننا نستطيع قياس انتاجية آلة أو سرعتها أو نكهة سلعة أو حجمها.

ان مصطلح الرقم في تعريف القياس يفرض بعض القيود على نوع المعالجة الرقمية الجائزة. فتستخدم الأرقام كرموز لبقاء نموذج للخصائص المرغوب بها في النظام العملي. ان طبيعة العلاقات التي تظهر في النظام العملي هي التي تحدد نوع المعالجة الرقمية التي تكون صالحة في النظام النظري. دعنا الآن نقوم بشرح خصائص نظام الأرقام حتى نتمكن من الحصول على فهم افضل لهذا الموضوع.

- خصائص نظام الأرقام Number System Characleristics

يتمتع نظام الأرقام (٠، ١، ٢، ٣، ٤، ٥، ٦، ٧، ٨، ٩) الذي قد تعلمناه من الدراسات الأساسية بأربعة خصائص أساسية هي :

١ - يعتبر كل رقم من هذه الأرقام مزيد، كما ان عدد هذه الأرقام هو عشرة أرقام.

٢ - ان ترتيب تسلسل الأرقام تمّ بالاتفاق أو التعارف. مثال (٢) أكبر من (١) و (١) أكبر من (صفر) .

٣ ـ يمكن تعريف أو تحديد فروقات متساوية. مثال ٤ – ٢ = ٨ – ٦، ٩ – ٥ = ٥ – ١ .

٤ ـ يمكن تعريف أو تحديد نسب متساوية. مثال ٦ ÷ ٣ = ٤ ÷ ٢ .

ان معالجة الأرقام باستخدام الاحصاء او الرياضيات يتضمن خاصية أو أكثر من خواص نظام الأرقام. وغالباً من يحاول الباحث جدياً استخدام أكثر من الخواص الرقمية التي يتطلبها النظام العملي خلال عملية تحليل البيانات. ولكن المشكلة تظهر في تحديد عدد الخصائص التي تظهر في الظاهرة الادارية قيد الدراسة ومن ثم تحديد أفضل خاصية من خواص نظام الأرقام التي يمكن استخدامها في عملية التحليل لمعالجة الأرقام.

ـ أنواع المقاييس Types of Scales

لقد تم تصنيف المقاييس حسب الخصائص الأربعة لنظام الأرقام. وهذه المقاييس هي : مقاييس اسمية، مقاييس ترتيبية، مقاييس فترات، مقاييس نسبة. ولقد تم تلخيص خصائص هذه المقاييس في جدول (٦ ـ ١). وفيما يلي شرح لكل ميزان أو مقياس من هذه المقاييس.

جدول (٦ ـ ١)

خصائص موازين القياس

الميزان	نظام الأرقام	الظاهرة الادارية	الاحصاء المستخدم*
الاسمي	التعريف المتميز للأرقام (٠، ١، ٢، ٣، ... ، ٩)	الانواع ذكر ـ انثى انواع المخازن مناطق البيع	النسب المنوال اختبار ذو الحدين اختبار X^{2}
الترتيبي	ترتيب الأعداد (٠ < ١ < ٢ < ... < ٩)	اتجاهات تفضيل الطبقة الاجتماعية الوظيفة	الوسيط ارتباط الرتب
الفترة	تساوي الفروق (٣ – ١ = ٩ – ٧)	اتجاهات آراء	المدى الوسط الانحراف المعياري
النسبة	النسب المتساوية ($\frac{٢}{٤} = \frac{٣}{٦}$)	العمر التكاليف عدد المستهلكين	الوسط الجبري معامل التغير

* كل الاحصائيات المناسبة للمقياس الاسمي تعتبر مناسبة للمقاييس الأعلى. وكذلك نفس الشيء بالنسبة للترتيب والفترة.

١ ـ الميزان الاسمي

ان الميزان الاسمي هو الميزان الذي يستخدم فيه الأرقام لتحديد أو ترتيب الأشياء أو الأحداث. واكثر الأمثلة شيوعاً هو استخدام الأرقام لتحديد لاعبو فرق الألعاب الرياضية مثل كرة القدم والطائرة ... الخ. فاذا كان الرقم (١٠) يعني اللاعب قلب الهجوم للفريق أ والرقم (٩) يعني اللاعب قلب الهجوم للفريق ب. فان الأرقام المستخدمة في هذا الحالة هي فقط لتحديد اللاعبين. وتفترض الأرقام المساواة في هذه الحالة بالنسبة لخصائص اللاعبين. لهذا فاننا لا نستطيع القول بان لاعب فريق أ هو افضل من لاعب فريق ب، بل تستخدم لتمييز اللاعبين.

ويستخدم القياس الاسمي لأقل اشكال القياس، وتسمية للتصنيف والتحديد. هناك بعض القيود المفروضة على عملية تعيين الأرقام للحوادث. ان قانون تعيين الأرقام ببساطة هو عدم اعطاء نفس الرقم لأكثر من حدث أو عدم اعطاء اكثر من رقم لنفس الحدث. فمثلاً يمكن تغيير رقم لاعب الهجوم لفريق أ من (١٠) الى (١٢) أو أي رقم آخر. ولكن لا نستطيع اعطاء اكثر من لاعب من نفس الفريق لنفس الرقم حتى نتمكن من استخدام الأرقام في تحديد هوية اللاعب. ولا نستطيع القول بان الرقم (١٠) هو افضل لوصف لاعب الهجوم من الرقم (١٢). ولكن جميع الأرقام تستخدم لأغراض التوصيف أو التحديد وبشكل متكافىء.

ان أنواع التحاليل الاحصائية الملائمة لمقياس البيانات الاسمية تتضمن المنوال والنسبة المئوية واختبار ذو الحدين واختبار X^2 . ان المنوال هو الاختبار الوحيد الذي يمكن أن يستخدم من بين مقاييس النزعة المركزية.

٢ ـ الميزان الترتيبي

يعرف الميزان الترتيبي ترتيب العلاقات ما بين الحوادث. وتتضمن هذا الميزان خاصية ترتيب الأرقام من بين خصائص النظام الرقمي أو العددي. وعليه فان هذا الميزان

يقيس ما اذا كان احدث الحوادث له خصائص اكبر او اقل من احد الحوادث الأخرى. ولكن لا يزود هذا المقياس حجم الزيادة او النقصان في الخصائص المختلفة التي تتمتع بها الحوادث المختلفة.

ان القانون الذي يتضمنه المقياس الترتيبي هو ان مجموعة الأرقام المستخدمة في التعيين تحافظ على ترتيب العلاقات التي تظهر في النظام العملي. وهذا يعني تعيين الأرقام التي تؤدي الى سلسلة من الأرقام المناسبة والتي تحافظ على ترتيب علاقات التفضيل للحدث. ولكن اي مجموعة من الأرقام التي يجب استخدامها في عملية التعيين؟ ان هناك مجموعات كبيرة من الأرقام يمكن تعيينها، ولكن القيد الوحيد هو تعيين الأرقام بطريقة بحيث تظهر العلاقات الموجودة في الحدث بشكل واضح. فعلى سبيل المثال يمكن استخدام الرقم (١) للتعبير عن ترتيب التفضيل الأول والرقم (٢) عن ترتيب الفضل الثاني وهكذا.

ان انواع التحاليل الاحصائية المناسبة لمقياس البيانات الترتيبية تتضمن الوسط وارتباط الرتب.

ويتضمن المقياس الترتيبي قطاع مهم من البيانات الادارية. ويتم تجميع معظم هذه البيانات عن طريق الأسئلة للأشخاص. ومن الأمثلة على ذلك قياس الاتجاهات والآراء والتفضيل. هذا بالاضافة الى ان كثير من خصائص البائعين او المشترين او عمال الانتاج او اي قسم آخر تتضمن خصائص ترتيبية مثل الوضع الاجتماعي ومستوى الوظيفة.

٣ ـ ميزان الفترة

يستخدم ميزان الفترة الأرقام لترتيب الحوادث بطريقة تكون فيها المسافات ما بين الأرقام تتوافق مع المسافة ما بين الحوادث بناءاً على الخصائص المستخدمة في القياس. ويمتلك ميزان الفترة جميع خصائص ميزان الرتب بالاضافة الى إحدى خصائص نظام

الارقام وهي خاصية المساواة في الفرق. ان افضل الامثلة المعروفة لميزان الفترة هو الميزان المئوي والميزان الفهرنهايتي لقياس الحرارة.

على الرغم من بيانات الاتجاهات والرغبات والتفضيل بيانات ترتيبية، الا ان استخدامها كبيانات فترة امر عادي في الابحاث الادارية. ولكن هناك عدم اتفاق بين الباحثين بالنسبة لمقدار الخطأ في القياس الناتج عن معالجة البيانات الترتيبية كبيانات فترة. لذلك يجب مقارنة قيمة هذا الخطأ بفوائد تحليل البيانات الناتج عن استخدام اختبارات احصائية متقدمة تصلح لتحليل بيانات الفترة. ومع هذا فان الشخص الباحث هو المسؤول عن تحديد ما يلي :

١ ـ قرب العلاقات التي تظهر في الظاهرة الادارية قيد الدراسة من ميزان الفترة.

٢ ـ مدى ملائمة معالجة البيانات في مقياس الفترة.

٤ ـ مقياس النسبة

يتمتع مقياس النسبة بجميع خصائص مقياس الفترة بالاضافة الى وجود خاصية وجود نقطة الصفر المطلق. ويتم تحديد رقم واحد فقط اعتياطياً باستخدام مقياس النسب، وبعد ذلك يتم اكمال تحديد الأرقام. ويتضمن مقياس النسب وجود نسب متساوية لقيم الميزان او المقياس تكافيء النسب المتساوية في خصائص الظاهرة الادارية قيد الدراسة.

ويوجد هناك كثير من الظواهر الادارية التي تمتلك خصائص مقياس النسبة. وتتضمن هذه الظواهر حجم المبيعات والعمر وعدد المستهلكين وعدد وحوادث الانتاج. كما يمكن استخدام جميع الاختبارات الاحصائية في تحليل بيانات ميزان النسبة.

أسئلة

س١ : ما هو المقصود بالقياس ؟

س٢ : ما هو الهدف من عملية القياس ؟

س٣ : لماذا يعتبر القياس صعب في الأبحاث الادارية ؟

س٤ : أذكر خصائص نظام الأرقام ؟

س٥ : قارن ما بين موازين القياس الأربعة ؟

الفصل السابع

طرق جمع البيانات

Data Collection Methods

طرق جمع البيانات
Data Collection Methods

- مقدمة

لقد تم التركيز في الفصول السابقة على دور البحث الاستكشافي في كل من عملية تشكيل مشاكل الدراسة واختيار منهجية البحث الملائمة في معالجة البحث قيد الدراسة. بمعنى آخر فقد تم التركيز على وجوب وضع مشاكل الدراسة بشكل واضح وكذلك تحديد المنهجية قبل البدء في مشروع البحث النهائي.

وعليه يتوجب على الشخص الباحث بعد تحديد الحاجة من البحث النهائي أن يحدد أهداف الدراسة بالاضافة الى تحديد المعلومات المتخصصة المطلوبة. لذلك فان الخطوة التي تتبع ذلك هو تحديد ما اذا كانت البيانات المرغوب بها متوفرة من مصادر البيانات الثانوية. ويمكن لمصادر البيانات الثانوية ان داخلية أو خارجية. ولكن غالباً لا تكفي البيانات الثانوية لتحقيق أهداف الدراسة قيد البحث. لذلك يتوجب على الشخص الباحث العمل على تشكيل تصميم للبحث يعتمد على مصادر البيانات الأولية. لذلك سوف يركز هذا الفصل على دراسة مصادر البيانات الأولية.

يعتبر الشخص المستجوب مصدراً رئيسياً للبيانات الادارية. ويوجد هناك طريقتان لجمع البيانات من المستجوب هي :

١ - الاتصال.

٢ - الملاحظة.

تتطلب طريقة الاتصال المشاركة الفعالة من المستجوب في تزويد البيانات

المرغوب بها، بينما تتضمن طريقة الملاحظة تسجيلاً لسلوك الشخص المستجوب.

وقبل الشروع بشرح الحالات المتعلقة بجمع البيانات، دعنا نستعرض أنواع البيانات التي يمكن تجميعها من الشخص المستجوب.

ـ أنواع البيانات التي يمكن تجميعها من المستجوب

ان اتخاذ القرارات الادارية غالباً ما يهتم بالقيام بأعماله في الوقت الحاضر حتى تتمكن من تحقيق اهداف مستقبلية. ومن هذا المنطلق فانه يمكن النظر الى الأبحاث الادارية كطريقة تنبؤ بحيث تكون مصمّمة لتسهيل عملية التوقع للظواهر الادارية. ان البيانات التي يمكن تجميعها من الشخص المستجوب لاستخدامها في تنبؤ الظواهر الادارية هي :

(١) سلوك ماضي. (٢) اتجاهات. (٣) خصائص الشخص المستجوب.

١ ـ السلوك الماضي.

لقد اثبتت الشواهد باتساع استخدام السلوك الماضي للشخص المستجوب في عملية توقع السلوك في المستقبل. وعليه فاننا كثير ما نستخدم في حياتنا اليومية شواهد السلوك الماضي في توقع سلوك مستقبلي لصديق أو قرابة. وبنفس المنطق فانه بالامكان للبحث الاداري من تجميع سلوك ماضي عن الشخص المستجوب بالنسبة لعدد الوحدات المنتجة أو الشراء أو باستخدام بعض السلع في توقع السلوك في المستقبل. لذلك يوجد هناك أكثر من محور يمكن استخدامه لفهم السلوك الماضي. وعليه يتوجب على الشخص الباحث أن يكون حساس بالنسبة للمحاور السلوكية الرئيسية المناسبة لتوقع السلوك المستقبلي عند تحديد البيانات المطلوبة لتحقيق حاجات الدراسة من المعلومات.

٢ ـ الاتجاهات.

تعتبر الاتجاهات ضرورية في دراسة معظم الظواهر الادارية بسبب الافتراض الموجود بالنسبة للعلاقة ما بين الاتجاهات والسلوك. فيمكن استخدام البيانات ذات العلاقة بالاتجاهات في تطوير استراتيجية أو تقييم برنامج تدريبي أو دعائي أو تقسيم السوق أو لتقييم طريقة من طرق الاشراف الادارية.

وبشكل عام تعتبر الاتجاهات ثلاثة مكونات رئيسية هي :

١ ـ الجانب الادراكي. ويهتم هذا الجانب بالاعتقاد الشخصي تجاه الظاهرة قيد الاهتمام. مثل سرعة آلة أو تعمير سلعة.

٢ ـ الجانب العاطفي. ويهتم هذا الجانب بالشعور الشخصي تجاه الظاهرة قيد الدراسة. مثل جيد أو سيء.

٣ ـ الجانب السلوكي. ويهتم هذا الجانب بدرجة الاستعداد الشخصي للاستجابة سلوكياً تجاه الظاهرة قيد الدراسة.

٣ ـ خصائص الشخص المستجوب.

ان الخصائص الشخصية تقوم بوصف المستجوب بناء على مجموعة من المتغيرات ذات الاهتمام. وتتضمن هذه المتغيرات كل من الخصائص الديمغرافية والسيكولوجية. وقد أثبتت الدراسات وجودعلاقة ما بين هذه المتغيرات والسلوك الشرائي. هذا بالاضافة الى استخدام بعض المتغيرات مثل العمر والجنس والحالة الاجتماعية وحجم العائلة والدخل والوظيفة والمستوى التعليمي في تقسيم مجتمع الدراسة أو العينة الى طبقات أو مجموعات.

- طرق جمع البيانات.

يوجد هناك طريقتين أساسيتين في تجميع البيانات هما الاتصال والملاحظة. وسوف نقوم الآن بشرح طريقة الاتصال أولاً ومن ثم ننتقل الى طريقة الملاحظة.

١ - طريقة الاتصال Communication Method

تعتمد طريقة الاتصال في تجميع البيانات على عملية طرح الأسئلة لجمهور المستجوبين. انه من المنطق ان تسال المستجوبين مجموعة من الأسئلة اذا أردت معرفة أي طريقة من طرق الاشراف الادارية يفضلها عمال قسم ما في مؤسسة ما، أو معرفة مدى تأثير برنامج تدريبي معين على انجاز العاملين او ما هي أسباب التسوق في سوق معين. ويمكن طرح الأسئلة اما شفوياً أو تحريرياً وكذلك نفس الشيء بالنسبة للاستجابة. ويطلق على الأداة المستخدمة في جمع البيانات اما بالاستبيان او اسئلة المقابلة الشخصية أو المقابلة الهاتفية. ولكن تعتبر أداة الاستبيان هي الأداة السائدة من بين هذه الأدوات في عمل الأبحاث الادارية.

- فوائد طريقة الاتصال

ان الفائدة الرئيسية التي تتمتع بها طريقة الاتصال هي التعدد. والمقصود بالتعدد هو مقدرة الأداة في جمع بيانات على مدى واسع من المعلومات المطلوبة. ان معظم مشاكل القرارات الادارية ذات علاقة مع الانسان. لهذا فان المعلومات المطلوبة لمشكلة القرار الاداري تركز على السلوك الماضي للانسان والاتجاهات بالاضافة الى الخصائص. ويمكن لطريقة الاتصال من تجميع البيانات عن هذه المجالات الثلاثة المختلفة.

ان الفائدة الأخرى ذات العلاقة بطريقة الأسئلة هي السرعة في الحصول على البيانات وقلة التكاليف مقارنة مع طريقة الملاحظة. ان التداخل والارتباط ما بين فائدة السرعة والتكلفة عالية. ان طريقة الاتصال طريقة سريعة في الحصول على بيانات مقارنة

مع طريقة الملاحظة بالاضافة الى مقدرتها في تزويد رقابة أفضل على عملية جمع البيانات. لذلك لا يتوجب على الشخص الباحث القيام بعملية التنبؤ لموعد أو مكان السلوك الذي يحدث أو الانتظار لحين حدوثه. فعلى سبيل المثال تكون عملية البيانات عن برنامج تدريبي معين عن طريق طرح الأسئلة أسرع وأرخص من محاولة التوقع أو ملاحظة الأشخاص المتدربين.

ـ سلبيات طريقة الاتصال.

يوجد هناك بعض السلبيات الرئيسية لطريقة الاتصال، أهمها :

١ ـ عدم رغبة الشخص المستجوب في تزويد البيانات المرغوب بها. ويتم ذلك إما عن طريق رفض المستجوب اعطاء الوقت الكافي للمقابلة أو رفضه للاجابة على أسئلة معينة.

٢ ـ عدم مقدرة الشخص المستجوب في تزويد المعلومات المطلوبة. قد لا يتمكن الشخص المستجوب من استرجاع بعض الحقائق المتعلقة بالسؤال أو قد لا يكون عنده علم بالموضوع قيد الدراسة من الأصل.

٣ ـ تاثير عملية الأسئلة على الاستجابات. قد يتحيز المستجوب في بعض الأحيان في الاجابة على بعض الأسئلة اما لكي يعطي اجابة مقبولة اجتماعياً أو لإرضاء الشخص المقابل.

على الرغم من ان هذه السلبيات تقلل من صلاحية طريقة الاتصال بشكل ملموس، الا انه بالامكان مراقبة هذه السلبيات عن طريق التصميم المناسب لأداة جمع البيانات. وسوف نقوم بشرح المزيد عن ذلك في الفصل اللاحق والذي يتعلق بتصميم أدوات جمع البيانات.

ـ أدوات الاتصال.

يمكن تقسيم طرق جمع البيانات حسب الأدوات المستخدمة في جمع البيانات الى ما يلي :

(١) المقابلة الشخصية. (٢) المقابلة الهاتفية. (٣) المقابلة البريدية.

١ ـ المقابلة الشخصية

تتكون المقابلة الشخصية من الشخص المقابل الذي يقوم بطرح مجموعة من الأسئلة لمستجوب واحد أو أكثر وجهاً لوجه. ان مهمة الشخص المقابل هي للاحتكاك مع الشخص أو مجموعة الأشخاص المستجوبين لطرح مجموعة من الأسئلة ومن ثم تسجيل الاجابات. لذلك يجب طرح الأسئلة بصورة واضحة ومن ثم تسجيل الاجابات بدقة. ويفضل ان يتم تسجيل الاجابات خلال المقابلة.

يمكن ان تؤدي عملية المقابلة الشخصية الى التحيز في اجابات الشخص المستجوب اما لإرضاء الشخص الباحث أو لإذهاله. وعليه فان هذا التحيز سوف يؤدي الى عدم صلاحية نتائج البحث في حل المشكلة قيد البحث.

٢ ـ المقابلة الهاتفية

تتكون المقابلة الهاتفية من الشخص المقابل الذي يقوم بطرح مجموعة من الأسئلة لمستجوب واحد أو أكثر بواسطة الهاتف بدلاً من الاتصال الشخصي المباشر. تستخدم المقابلة الهاتفية بشكل واسع لمقدرتها في تجميع نطاق واسع من البيانات بالاضافة الى الاقتصاد في اجراءات جمع البيانات.

ان التحيز الناتج عن التداخل الاجتماعي ما بين الشخص المقابل والشخص المستجوب في الطريقة الهاتفية يكون قليل مقارنة مع المقابلة الشخصية. ان السلبية الأساسية في المقابلة الهاتفية تعود الى محدودية كمية المعلومات التي يمكن الحصول

عليها. هذا بالاضافة الى التحيز الناتج اما عن عدم السماع الكامل لاجابات الأشخاص المستجوبين أو لعدم وجود قائمة كاملة بأسماء مجتمع الدراسة.

٣ ـ المقابلة البريدية

تتكون المقابلة البريدية من استبيان يرسل الى الشخص المستجوب ومن ثم اعادة الاستبيان بعد تعبئته الى الشخص الباحث أو مؤسسة الأبحاث. وتتميز المقابلة البريدية بالمرونة في التطبيق وقلة في التكاليف وتخلو من التحيز الناتج عن التداخل الاجتماعي ما بين الشخص الباحث والشخص المستجوب. ان السلبية الرئيسية ذات العلاقة بالمقابلة البريدية هي مشكلة الخطأ الناتج عن عدم الاستجابة.

ويوجد هناك أكثر من طريقة لتوزيع الاستبيان وتجميعه. فمن الممكن ان يتم توزيعه وتجميعه عن طريق الانسان بدلاً من البريد. ومن الممكن ان يوزع الاستبيان بالمجلات والصحف.

ـ المعايير المستخدمة في اختيار وسيلة الاتصال.

يوجد هناك العديد من المعايير المناسبة لتقييم وسيلة الاتصال التي تحقق حاجة البحث قيد الدراسة. وهذه المعايير هي :

(١) التعدّد. (٢) التكاليف. (٣) الوقت. (٤) مراقبة العينة.

(٥) كمية البيانات. (٦) نوعية البيانات. (٧) نسبة الاستجابة.

(١) التعدّد

يشير التعدد الى مقدرة وسيلة الاتصال في تسخير أو اخضاع عملية جمع البيانات لخدمة الحاجات الخاصة بالدراسة قيد البحث. وأكثر الطرق مقدرة في عمل ذلك من بين هذه الطرق الثلاثة هي المقابلة الشخصية. بينما مقدرة طريقة الهاتف في

تسخير عملية جمع البيانات لخدمة حاجات البحث الخاصة أقل من طريقة المقابلة الشخصية. كما تعتبر طريقة المقابلة البريدية من أقل الطرق مقدرة في عمل ذلك.

ان المقابلة الشخصية ما بين الشخص الباحث والمستجوب تمكن الشخص الباحث من توضيح وشرح الأسئلة المعقدة وعرض بعض الحالات التي تساعد في فهم أسئلة المقابلة للشخص المستجوب كجزء من عملية البحث. أما بالنسبة للمقابلة الهاتفية فلا يتم هناك مقابلة شخصية مما يجعل من الصعب تضمين المقابلة بأسئلة معقدة والتي تتطلب اجابات متخصصة للأسئلة المفتوحة. أما اذا كان تصميم البحث والتعليمات المزودة تجعل الاجابة على الأسئلة سهلة، فان غالباً ما يؤدي هذا الى الغاء دور الشخص المقابل وبالتالي تكون المقابلة البريدية من افضل الطرق لاتمام دراسة البحث.

يتوجب على الشخص الباحث تحديد درجة التقلّب التي يحتاجها البحث قيد الدراسة حتى يتمكن من اختيار افضل ادوات الاتصال في تحقيق حاجات الدراسة. وغالباً ما تكون درجة التقلب في الأبحاث عملياً قليلة لدرجة تسمح باستخدام طريقة المقابلة البريدية.

(٢) التكاليف

ان عدد ساعات العمل المطلوبة لكل اداة من أدوات الاتصال غالباً ما تحدد نسبة التكاليف لكل وسيلة. وتتضمن تكلفة العمل اجرة الشخص الباحث وتكاليف الاشراف المتعلقة بمراقبة نوعية عملية جمع البيانات. وتعتبر طريقة المقابلة الشخصية من أكثر طرق تكلفة في عملية جمع البيانات. اما بالنسبة للطريقة الهاتفية فتأتي بعد المقابلة الشخصية أما بالنسبة للطريقة البريدية فتعتبر أقل طرق تكلفة بالنسبة لجمع البيانات.

(٣) الوقت

تعتبر طريقة المقابلة الهاتفية من أسرع الطرق في جمع البيانات المطلوبة خاصة اذا كانت أسئلة الاستبيان أو المقابلة قليلة . ان طريقة المقابلة الشخصية تحتاج الى وقت للانتقال ما بين الأشخاص المستجوبين بالاضافة الى وقت المقابلة نفسه .

ويمكن تقليل الوقت الكلي اللازم لاتمام مشروع البحث عن طريق زيادة عدد المقابلين الذين يعملون في الدراسة عند استخدام المقابلة الهاتفية أو الشخصية . أما بالنسبة للمقابلة البريدية فانه من الصعب جداً تقليل الوقت الكلي اللازم للدراسة . فلا يستطيع الشخص الباحث عمل اي شيء من شأنه ان يسرّع الاجابة في تعبئة الاستبيان ومن ثم اعادته الى شخصه بعد توزيع الاستبيانات بالبريد .

(٤) مراقبة العينة

تشير مراقبة العينة الى قدرة وسيلة الاتصال في الوصول الى جميع الوحدات المدرجة في خطة العينة بكفاءة وفعالية . ان الاختلاف ما بين وسائل الاتصال بهذا الخصوص واضح بشكل كبير .

تزود المقابلة الشخصية أفضل درجة في مراقبة العينة مقارنة مع المقابلة الهاتفية أو البريدية . فاذا كان الاطار العيني غير موجود ، فان اجراءات اختيار العينة تعتمد بشكل كبير على المقابلات الشخصية . ان العمل في المقابلات الشخصية يمكن الشخص الباحث من مراقبة الوحدات العينية التي تمت مقابلتها ودرجة التعاون وجوانب عديدة أخرى في عملية جمع البيانات .

تعتمد المقابلة الهاتفية بشكل كبير على اطار العينة . وغالباً ما تستخدم الأدلة الهاتفية في هذه الحالة كاطار للعينة لاختيار عناصر العينة عن طريق اجراءات الاختيار الاحتمالية . وغالباً ما تعتبر الأدلة الهاتفية اطار ضعيف للعينة لعدم احتواءها على جميع أفراد مجتمع الدراسة للأسباب التالية :

١ ـ عدم توفر خدمة الهاتف لكل فرد من افراد المجتمع.

٢ ـ غالباً ما تكون الأدلة قديمة ولا تحتوي على أسماء المنتفعين الجدد من الخدمة الهاتفية.

٣ ـ بعض أسماء الأشخاص أصحاب الخدمة الهاتفية غير مدرجة بالدليل.

وبما ان الأدلة الهاتفية غير ممثلة لمجتمع الدراسة فان المقابلة الهاتفية تسمح بمراقبة محدودة على العينة.

تتطلب المقابلة البريدية قائمة بجميع عناصر مجتمع الدراسة قيد البحث كما هو الحال في المقابلة الهاتفية. ويتكون اطار العينة بالنسبة للمقابلة البريدية من الأسماء والعناوين. وفي بعض الحالات يصعب تحديد عناوين جميع عناصر العينة مما يؤدي الى تقليل درجة المراقبة والتي تسمح بها المقابلة البريدية على العينة.

(٥) كمية البيانات

ان أكثر وسيلة من بين الوسائل الثلاثة مقدرة في تجميع أكبر كمية من البيانات هي وسيلة المقابلة الشخصية ثم المقابلة الهاتفية وأخيراً المقابلة البريدية. كمية البيانات التي يمكن تجميعها باستخدام هذه الوسائل المختلفة كبيرة اذا كان هناك علاقة عاطفية ما بين الأشخاص المستجوبين وموضوع البحث. ولكن في الوضع الطبيعي تعتبر طريقة المقابلة الشخصية من أكثر الطرق مقدرة في جمع البيانات.

ان الايجابية الرئيسية للمقابلة الشخصية تنتج من العلاقة الاجتماعية ما بين الشخص المقابل والمستجوبين. وغالباً ما تعمل العلاقة الاجتماعية على تحفيز الشخص المستجوب في اعطاء وقت أكبر للمقابلة. بينما من السهل جداً أن يقوم الشخص المستجوب بتحديد أو انهاء وقت المقابلة في كل من طريقة المقابلة الهاتفية أو الطريقة البريدية.

ان الميّزة التي تتمتع بها كل من طريقة المقابلة الشخصية والهاتفية على الطريقة البريدية هي قلّة الجهد المطلوب من قبل الشخص المستجوب في عملية جمع البيانات. فالشخص المقابل هو الذي يطرح الأسئلة ومن ثم يتحسّس الاجابات وبعد ذلك يقوم بتسجيلها.

(٦) نوعية البيانات

تشير نوعية البيانات الى الدرجة التي تكون فيها البيانات خالية من التحيز الشخصي الناتج عن استخدام وسيلة الاتصال. فاذا كان موضوع البحث غير عاطفي وتصميم الأسئلة مناسب، فاننا نتوقع الحصول على بيانات ذات نوعية بغض النظر عن الأداة المستخدمة في عملية جمع البيانات.

وقد اثبتت خبرة الباحثين وجود فروق فعلية ما بين الوسائل الثلاثة في جمع البيانات اذا كانت المقابلة تحتوي على اسئلة حساسة او محرجة مثل السؤال عن رصيد البنك او الدخل او بعض السلوك الخاص. فاذا كانت الأسئلة حساسة فتعتبر طريقة المقابلة البريدية من افضل الطرق لجمع بيانات ذات نوعية جيدة مقارنة مع المقابلة الشخصية. وغالباً ما تقع المقابلة الهاتفية في مكان ما بين هاتين الطريقتين.

ان المصدر الآخر للتحيز ناتج عن ارتباك او عدم فهم الشخص المستجوب لأسئلة المقابلة. فاذا اراد الشخص المستجوب اي توضيح عن الأسئلة من الشخص الباحث الغائب فان غالباً ما يؤدي الى جمع بيانات غير دقيقة نتيجة الارباك. كما ان الطريقة الهاتفية تزود بيانات متحيزة ناتجة عن الارباك أكثر من المقابلة الشخصية لغياب الوجود الجسمي للشخص الباحث.

ان مصدر التحيز في الطريقة البريدية ينتج عن طريق قراءة الشخص المستجوب للأسئلة قبل البدء بالاجابة او تغيير في بعض الاجابات للأسئلة التي تقع في بداية الاستبيان بسبب تأثير الأسئلة التي تقع في نهاية الاستبيان. بينما نجد خلو كل من

الطريقتين المقابلة الشخصية والهاتفية من هذا النوع من التحيز.

كما ان كل من طريقتي المقابلة الشخصية والهاتفية قادرة على تجميع بيانات عن الأحداث وقت وقوعها. وهذا يقلل من التحيّز الناتج عن عدم مقدرة الشخص المستجوب في استرجاع الأحداث بدقة.

ان آخر اعتبار بالنسبة لنوعية البيانات المجمّعة يهتم بغش الشخص المقابل. ان الغش عملية سهلة في المقابلة الشخصية مقارنة مع الطريقة الهاتفية اذا تم تسجيلها مباشرة أو الطريقة البريدية لعدم وجود الشخص الباحث خلال عملية المقابلة. ولكن يمكن تقليل تأثير هذا المصدر بالنسبة للتحيز اذا ما تم تصميم المقابلة الشخصية بطريقة فعالة.

(٧) نسبة الاستجابة (الردّيات).

تشير نسبة الاستجابة الى نسبة الأشخاص التي تمت مقابلتهم فعلاً الى العدد الذي يرغب الباحث مقابلته. فاذا كانت نسبة الاستجابة قليلة فان ذلك سوف يؤدي الى وجود أخطاء عدم الاستجابة بشكل كبير مما يؤدي الى عدم صلاحية نتائج الدراسة. وتمثل أخطاء عدم الاستجابة الفرق ما بين الأشخاص الذين استجابوا للدراسة والأشخاص الذين لم يستجيبوا. ويعتبر الخطأ الناتج عن عدم الاستجابة من أكثر الأخطاء التي تواجه الباحث خطورة. ويتزايد احتمال تحقيق هذا النوع من الأخطاء كلما قلت نسبة الاستجابة.

ويوجد هناك مصدرين لعدم الاستجابة هما : (١) اذا كان الشخص المستجوب غير موجود في المنزل. (٢) رفض الشخص المستجوب للاستجابة. ان تأثير المصدر يكون كبيراً بالنسبة لطريقة المقابلة الشخصية أو الهاتفية، بينما يكون تأثيره على الطريقة البريدية محدد. وبما ان احتمال أن يجيب الشخص المستجوب على الهاتف أكبر من احتمال فتح باب المنزل في حالة وجود شخص غريب، فان تأثير هذا المصدر

على الطريقة الهاتفية يكون أقل مقارنة مع المقابلة الشخصية.

ان إحدى الطرق الرئيسية المستخدمة لتقليل تأثير رفض الشخص المستجوب للتعاون مع الشخص الباحث هي عن طريق تطوير علاقات شخصية مع أكثر المستجوبين تأثيراً على موضوع البحث. ان مشكلة عدم الاستجابة في الطريقة البريدية ناتجة عن عدم رغبة الشخص المستجوب أحياناً في موضوع البحث قيد الدراسة. وللتقليل من تأثير هذه المشكلة، فقد اقترح روبن خمسة اجراءات هي :

(١) تزويد الشخص المستجوب برسالة قبل الاستبانة.

(٢) ارسال الاستبيان مع رسالة تغطية.

(٣) ارسال رسالة للمتابعة.

(٤) ارسال استبيان ثاني.

(٥) بعث رسالة متابعة ثالثة.

ـ اختيار وسيلة الاتصال

ما هي أدوات الاتصال التي يجب اختيارها لتنفيذ البحث قيد الدراسة؟ ان الاجابة على هذا السؤال هي اختيار الوسيلة التي تضمن جمع بيانات تعمل على تحقيق المعلومات المطلوبة باقل وقت وتكلفة. يجب ملاحظة هنا ان هذه الأدوات الثلاثة في جمع البيانات غير مانعة لبعضهما البعض. بمعنى آخر يمكن للباحث استخدام أكثر من وسيلة لجمع البيانات عند تصميم البحث. وهذا سوف يسمح للشخص الباحث من تصميم دراسته بشكل يستطيع من خلاله الاستفادة من نقاط القوة لكل أداة من هذه الأدوات المستخدمة في تجميع البيانات.

٢ ـ طريقة الملاحظة

تتضمن الملاحظة تسجيل سلوك المستجوب. وعليه فالملاحظة هي معرفة وادراك سلوك الانسان او الهدف أو الحدث وثم القيام بتسجيله. وتستخدم طريقة المراقبة غير الرسمية بشكل كبير من متخذي القرارات. وتتمثل الخطورة في عمل استنتاجات عن طريق المراقبة غير الرسمية بالوجود الكبير للأخطاء الناتجة عن العينة وغير العينة. لذلك ان طرق المراقبة الرسمية يأخذ بعين الاعتبار مراقبة هذه الأخطاء لتزويد بيانات صالحة لاتخاذ القرارات. ومن النادر جداً أن تجد تصميم بحث يعتمد كلياً على طريقة الملاحظة. وسوف نقوم في الجزء بشرح الطرق المختلفة والمستخدمة في الملاحظة الرسمية. ولكن قبل ذلك دعنا نتعرف على ايجابيات وسلبيات هذه الطريقة.

ـ ايجابيات طريقة الملاحظة.

يوجد هناك عدة ايجابيات لطريقة الملاحظة مقارنة مع طريقة الأسئلة. وهذه الايجابيات هي :

١ ـ لا تعتمد طريقة الملاحظة في جمع البيانات المرغوب بها على استعداد الشخص المستجوب في ذلك.

٢ ـ تقليل أو الغاء التحيز الناتج اما بسبب الشخص المقابل أو عملية المقابلة.

٣ ـ عدم امكانية تجميع بعض البيانات الا عن طريق الملاحظة.

ومثال على ذلك : جميع الأنماط السلوكية التي لا يدركها الشخص المستجوب.

ـ سلبيات طريقة الملاحظة.

يوجد هناك سلبيتان رئيسيتان لطريقة الملاحظة مما يحدد استخدام هذه الطريقة. وهاتان السلبيتان هما :

١ ـ عدم مقدرة هذه الطريقة في ملاحظة بعض الأشياء كالادراك والاعتقاد والشعور والتفضيل.

٢ ـ اذا كان الوقت المتاح للملاحظة والمال محدودين، فانه من الصعب جداً ملاحظة أي نمط من الأنماطالسلوكية للفرد الملاحظ الا اذا تكرر السلوك بشكل مستمر. وهذا سوف يحدد استخدام طريقة الملاحظة لمجموعة من الحالات الفريدة.

ـ تصنيف طرق الملاحظة.

يمكن تصنيف طرق الملاحظة بخمسة طرق مختلفة هي :

(١) الملاحظة الطبيعية ضد الاستنباطية.

(٢) الملاحظة التنكرية ضد عدم التننكرية.

(٣) الملاحظة المنظمة ضد غير المنظمة.

(٤) الملاحظة الانسانية ضد الآلية.

(١) الملاحظة الطبيعية ضد الاستنباطية.

تتضمن الملاحظة الطبيعية ملاحظة السلوك وكما يحدث في البيئة الطبيعية مثل التسوق في بقالة تموينية. اما الملاحظة الاستنبطاية فتتضمن تشكيل بيئة مصطنعة ومن ثم ملاحظة انماط السلوك لفرد يوضع في هذه البيئة، مثل جعل الناس يتسوقون في نموذج لبقالة تموينية.

ان الايجابية الرئيسية للملاحظة في البيئة الطبيعية تتمثل في مقدرة السلوك المعروض في هذه البيئة على تمثيل انماط السلوك الحقيقية. ولكن يجب على الباحث أن يزن هذه الايجابية مع الزيادة المتحققة بالتكاليف نتيجة انتظار وقوع السلوك الطبيعي، بالاضافة الى صعوبة قياس السلوك في البيئة الطبيعية.

(٢) الملاحظة التنكرية ضد اللاتنكرية.

تشير التنكرية الى ما اذا كان الشخص المُلاحِظ يدرك انه تحت المراقبة أم لا. ويجب أن يكون دور المُلاحِظ تنكرياً في جميع الحالات التي يكون فيها سلوك الانسان مختلف عن السلوك الطبيعي اذا ما علم انه تحت المراقبة. وهناك أكثر من طريقة يمكن استخدامها لاخفاء عملية المراقبة مثل الكاميرا الخفية.

وهناك اختلاف ما بين الباحثين حول درجة تاثير وجود الشخص المُراقب على الأنماط السلوكية للناس تحت المراقبة. فمنهم من يقول بان التاثير قليل ومنهم من يقول بأن التاثير كبير ويؤدي الى تحيز خطير في أنماط السلوك الملاحظة.

(٣) الملاحظة المنظمة ضد غير المنظمة.

تعتبر الملاحظة المنظمة مناسبة عندما تكون مشكلة القرار محددة وبوضح والمعلومات المطلوبة تسمح في تحديد انماط السلوك المراد مراقبتها وقياسها. اما بالنسبة للملاحظة غير المنظمة فتعتبر مناسبة في الحالات التي تكون فيها مشكلة القرار غير محددة بعد، بحيث تكون الحاجة للمعلومات الكثيرة ملحة لتطوير وتشكيل فرضيات تساعد في تعريف المشكلة قيد الدراسة.

وتعتبر الملاحظة المنظمة أكثر ملائمة بالنسبة لدراسات الأبحاث النهائية. ان تنظيم الملاحظة يقلل من فرصة التحيز عند الشخص الذي يقوم بعملية الملاحظة ويزيد من صلاحية البيانات.

وتعتبر الملاحظة غير المنتظمة أكثر ملائمة بالنسبة لدراسات الأبحاث الاستكشافية. فالشخص الباحث (المراقب) هنا يكون حر بالنسبة لمراقبة الأنماط السلوكية التي تعتبر مناسبة لحالة القرار. وهذا يعني وجود فرصة كبيرة للتحيز من قبل الباحث (الملاحظ). لذلك يجب أن تعامل نتائج الدراسة كفرضيات مطلوب قياسها بتصميم بحث نهائي.

(٤) الملاحظة الانسانية ضد الآلية.

يفضل في بعض الحالات استبدال المراقب الانسان بشكل معين من المراقبة الآلية. وقد يكون السبب اما لزيادة الدقة في البيانات المجمعة أو لتخفيض التكاليف أو متطلبات لبعض المقاييس الخاصة. وتتضمن الوسائل الآلية التي تستخدم في المراقبة ما يلي :

(١) كاميرا الصور المتحركة.

(٢) كاميرا العين.

(٣) مقياس بؤبؤ العين.

تستخدم كاميرا الصور المتحركة لتسجيل السلوك الشرائي في متجر أو صيدلية وما الى ذلك. وهنا يقوم الملاحظ بتقييم فيلم الكاميرا وقياس السلوك المرغوب به. ان استخدام وسائل متعددة للمراقبة واعادة المشاهدة تسمح بقياس السلوك بدقة أكبر.

تستخدم كاميرا العين لقياس حركة العين. وتستخدم هذه الوسيلة لتحديد كيف يقرأ شخص المجلة أو الجريدة أو الدعاية أو اي شيء آخر مشابه. ويتم القياس حسب تسلسل وقوع الأشياء المراقبة والوقت الذي انقضى بالنظر الى الأجزاء المختلفة.

اما مقياس بؤبؤ العين فيستخدم لقياس التغيّر في قطر البؤبؤ للعين. ومن المفروض ان تعكس الزيادة في قطر العين التفاعل الايجابي بالنسبة للمؤثر الذي تمت ملاحظته.

أسئلة

س١ : أذكر أنواع البيانات الرئيسية التي يمكن تجميعها من المستجوب ؟

س٢ : أذكر ايجابيات وسلبيات طرق الاتصال ؟

س٣ : أذكر وسائل الاتصال المستخدمة في تجميع بيانات المستجوب، وما هي المعايير المستخدمة في المفاضلة بينهما ؟

س٤ : أذكر ايجابيات وسلبيات طرق الملاحظة ؟

الفصل الثامن

تصميم أشكال جمع البيانات

Designing Data Collection Forms

تصميم أشكال جمع البيانات
Designing Data Collection Forms

ـ مقدمة

لقد ركزت في الفصول السابقة على شرح البيانات بنوعيها الثانوية والأولية. اما في هذا الفصل فسوف يتم شرح الأمور المتعلقة بتصميم اشكال جمع البيانات من خلال التركيز على بناء الأشكال المناسبة في جمع البيانات للأبحاث النهائية. علماً بن تصميم اي بحث يتطلب طريقة جمع بيانات قادرة على تزويد بيانات صالحة ومناسبة للمشكلة قيد الدراسة.

تعتبر اشكال جمع البيانات المكون الرئيسي لمعظم دراسات الأبحاث. كما ان ادوات طريقة الاتصال ـ المقابلة الشخصية والهاتفية والبريدية ـ تعتمد على الاستبيان في تجميع البيانات. وبما ان دراسات الاتصال اكثر شيوعاً من دراسات الملاحظة، فانه سوف يتم التركيز في هذا الفصل بشكل رئيسي على تطوير الاستبيان وتخصيص القسم النهائي على اشكال الملاحظة.

ـ مكونات الاستبيان.

ان الاستبيان عبارة عن جدول منظم لجمع البيانات من جمهور مجتمع الدراسة. ووظيفة الاستبيان هو القياس. ويمكن استخدام الاستبيان في قياس الأمور التالية :

(١) سلوك ماضي.

(٢) اتجاهات.

(٣) خصائص شخصية.

يتكون الاستبيان النموذجي من خمسة أجزاء هي :

(١) بيانات معرّفة.

(٢) طلب مساعدة.

(٣) تعليمات.

(٤) المعلومات المطلوبة.

(٥) بيانات تصنيفية.

١ ـ البيانات المعرّفة.

غالباً ما تحتل البيانات المعرّفة الجزء الأول من الاستبيان. وغالباً ما تتضمن هذه البيانات اسم الشخص الباحث وعنوانه ورقم هاتفه. كما تتضمن هذه البيانات وقت المقابلة وتاريخها واسم الشخص المقابل.

٢ ـ طلب المساعدة.

يعتبر طلب المساعدة جملة افتتاحية مصممة لتشجيع أعضاء مجتمع الدراسة في مساعدة الباحث بالنسبة للمقابلة. وغالباً ما تحدد هذه الجملة الشخص المقابل أو المؤسسة المقابلة. بالاضافة الى توضيح الغرض من الدراسة وتحديد الوقت اللازم لاتمام الدراسة.

(٣) التعليمات.

تشير التعليمات الى الملاحظات الموجهة من الشخص المقابل أو الباحث الى أفراد مجتمع الدراسة عن كيفية استخدام الاستبيان. وغالباً ما تظهر هذه الملاحظات في

الاستبيان بشكل مباشر عند استخدام المقابلة البريدية. أما في الطريقة المقابلة الشخصية أو الهاتفية فغالباً ما تظهر التعليمات في ورقة مستقلة توضح الغرض من الدراسة واطار العينة وأي جانب آخر من جوانب عملية جمع البيانات. هذا بالاضافة الى احتواء الاستبيان الى بعض التعليمات الخاصة بالنسبة للاجابة على بعض الأسئلة. ومن الأمثلة على ذلك طرق قياس الاتجاهات.

(٤) المعلومات المطلوبة.

تشكل المعلومات المطلوبة الجزء الرئيسي من الاستبيان. ان الجزء اللاحق من هذا الفصل سوف يركز على تصميم هذا الجانب من الاستبيان.

(٥) بيانات تصنيفية.

تهتم البيانات التصنيفية بخصائص الفرد المستجوب. وغالباً ما تزود هذه البيانات مباشرة من الشخص المستجوب في حالة استخدام المقابلة البريدية. اما في المقابلة الشخصية او الهاتفية فغالباً ما يتم تزويد هذه البيانات عن طريق الشخص المقابل او الباحث. وغالباً ما يتم تجميع هذه البيانات في نهاية المقابلة. ولكن بعض اجراءات العينة تتطلب تجميع هذه البيانات في بداية المقابلة لتحديد ما اذا الفرد أو الشخص مؤهل لأن يكون جزء من اطار العينة.

ـ تصميم الاستبيان

يعتبر تصميم الاستبيان فن أكثر منه موضوع علمي يمكن دراسته. فلا يوجد هناك خطوات متسلسلة او مبادىء يمكن اتباعها بحيث تضمن بناء استبيان كفىء وفعال. ان تصميم الاستبيان عبارة عن مهارة يكتسبها الباحث من خلال الخبرة لا من قراءة خطوط استرشادية متسلسلة. والطريقة الوحيدة لتنمية هذه المهارة هي عن طريق

كتابة الاستبيان واستخدامه في سلسلة من المقابلات وتحليل نقاط الضعف ومن ثم اعادة كتابته. وهكذا.

ان الشيء الذي نعرفه عن تصميم الاستبيان يأتي من خبرة الباحثين المتخصصين في هذا الموضوع. وبناءاً على هذه الخبرة المتراكمة فاننا نستطيع الخروج بمجموعة من القوانين أو الخطوط الاسترشادية التي يمكن أن تكون مفيدة للباحثين الجدد في عملية تصميم الاستبيان. فبينما هذه القوانين ضرورية لتفادي الأخطاء الخطيرة، الا ان تنقية تصميم الاستبيان تأتي من الباحث الماهر. لذلك تعتمد نوعية الاستبيان على مهارة وحكم الشخص الباحث وفهمة للمعلومات المطلوبة.

وسوف يتم تنظيم شرح تصميم الاستبيان بسبعة خطوات. وسوف يتم عرض النقاط الاسترشادية في بناء الاستبيان عند كل مرحلة من هذه المراحل. وهذه الخطوات والموضحة في الشكل (٨ - ١) هي :

(١) مراجعة الاعتبارات الأساسية.

(٢) تحديد محتوى كل سؤال.

(٣) تحديد شكل الاستجابة.

(٤) تحديد كلمات كل سؤال.

(٥) تحديد ترتيب الأسئلة.

(٦) تحديد الخصائص المادية للاستبيان.

(٧) الاختبار الأولي والمراجعة والشكل النهائي للاستبيان.

شكل (٨ ـ ١)

خطوات تصميم الاستبيان

١ ـ الاعتبارات الأولية

↓

٢ ـ محتوى الأسئلة

↓

٣ ـ شكل الاستجابة

↓

٤ ـ كلمات السؤال

↓

٥ ـ تسلسل الأسئلة

↓

٦ ـ الخصائص المادية

↓

٧ ـ الاختبار الأولي والمراجعة والشكل النهائي

١ ـ مراجعة الاعتبارات الأساسية Review Preliminary Considerations

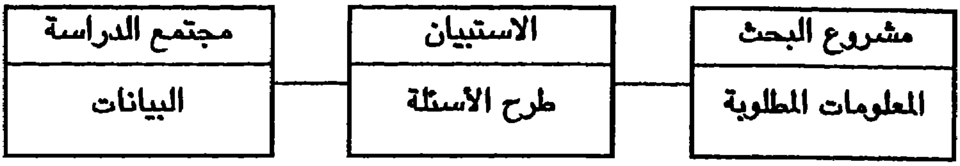
تركز الاعتبارات الأساسية عند تنفيذ البحث النهائي على طرق تطوير وسيلة ربط فعالة ما بين عملية اتخاذ القرار وعملية البحث. وتعتبر عملية تطوير أهداف الدراسة وتحديد المعلومات المطلوبة مركزاً لهذه العملية. وعليه يجب تصميم البحث بالاضافة الى بلورة وتخطيط عملية البحث قبل تصميم الاستبيان. بمعنى آخر ان تصميم الاستبيان يسبقه مجموعة من القرارات قبل البدء بتنفيذه مثل نوع تصميم البحث ومصادر البيانات ودور الاستبيان في عملية البحث. لذلك، فانه لمن الضروري جداً أن يمتلك الشخص الباحث صورة واضحة عن مجتمع الدراسة ومعرفة مفصلة عن اطار العينة. لأن تطوير الاستبيان يتأثر بشكل كبير بخصائص مجتمع الدراسة. فكلما زاد عدم تجانس افراد مجتمع الدراسة، كلما زادت صعوبة تصميم استبيان واحد بحيث يعتبر مناسباً لجميع افراد المجتمع. وعليه يجب أن يصمم الاستبيان بطريقة بحيث يكون فيها مفهوم لأقل شخص مقدرة على الفهم من بين مجموعة أفراد مجتمع الدراسة. كما يجب تحديد ميزان القياس ووسيلة الاتصال. هذا بالاضافة الى بلورة عملية اعداد البيانات ومراحل التحليل وطبيعية نتائج البحث. ان النواحي الفنية في تصميم الاستبيان يعتمد بشكل كبير على المراحل النهائية لعملية البحث.

كما يعتمد تصميم الاستبيان على علاقة الربط ما بين قائمة المعلومات المطلوبة والبيانات المراد تجميعها وبشكل مفصل. ان الاستبيان هو أداة ربط بين المعلومات المطلوبة والبيانات المراد تجميعها. والشكل (٨ ـ ٢) يوضح طبيعة هذه العلاقة.

كما يجب أن تكتب الأسئلة بتسلسل منطقي يعكس قائمة المعلومات المطلوبة. لذلك يجب عدم كتابة أي سؤال في الاستبيان الا اذا كانت له علاقة مع معلومة محددة مطلوبة. ولا يوجد في الحياة العملية نزعة قوية لتضمين أي سؤال لانه ممتع الا اذا كان له علاقة ربط قوية مع المعلومات المطلوبة. ان تضمين الاستبيان أسئلة غير ضرورية تزيد من تكاليف الدراسة وكذلك تزيد من الجهد المطلوب من المستجوب.

شكل (٨ ـ ٢)

ربط المعلومات المطلوبة مع البيانات المجمعة



٢ ـ تحديد محتوى الأسئلة Decide on Question Content

يتأثر محتوى أسئلة الاستبيان بمجموعة من الاعتبارات المتعلقة بالشخص المستجوب هي : (١) المقدرة على الاجابة بدقة . (٢) الاستعداد للاستجابة بدقة .

أ ـ المقدرة على الاجابة بدقة . Ability to Answer Accuratelly

يتوجب على الشخص الباحث ان يكون حساساً بالنسبة لمقدرة الشخص المستجوب في تزويد البيانات على فرض ان البيانات المرغوب بها ملائمة لمشكلة القرار . ويوجد هناك أنواع من البيانات لا يمكن تجميعها من الشخص المستجوب بدقة . ويمكن للبيانات غير الدقيقة أن تنتج من مصدرين هما :

(١) عدم معرفة الشخص المستجوب .

(٢) نسيان الشخص المستجوب .

١ ـ عدم معرفة الشخص المستجوب

ان كثير من الأسئلة التي تطرح علينا في الحياة اليومية قد لا نجد لها جواب . يمكن طرح مجموعة من الأسئلة على مجموعة من الأشخاص حول ظاهرة معينة بحيث لا يكون أحد هذه الأفراد مدركاً لهذه الظاهرة .

وقد اكتشف الباحثون بالتجربة ان هناك مجموعة من المستجوبين غالباً ما يقوموا بالاجابة على بعض الأسئلة على الرغم من عدم معرفتهم بموضوع هذه الأسئلة لعدم رغبتهم بالاعتراف بهذا النوع من العجز. لذلك فان هذا الوضع يمثل مصدر خطير من مصادر أخطاء القياس.

ان طريقة كتابة بعض الأسئلة تشجع الاجابة عليها من قبل المستجوبين لافتراضها وجوب معرفة الاجابة على الأسئلة. اعتبر السؤال الآتي: «ما هي نسبة الفائدة الحالية التي تتقاضاه على حساب التوفير؟ » ان هذا السؤال يتضمن وجوب معرفة الشخص المستجوب الاجابة على هذا السؤال. ان السؤال البديل عنه هو : «هل تعلم نسبة الفائدة الحالية التي تتقاضاه على حساب التوفير؟ » ان هذا السؤال يتضمن عدم معرفة بعض الناس لنسبة الفائدة الحالية مما يجعل اعتراف المستجوب بعدم معرفته بنسبة الفائدة أسهل.

٢ ـ نسيان الشخص المستجوب

ان كثير من الأسئلة التي تُطرح علينا في الحياة اليومية غالباً ما نمتلك الاجابة عليها في وقت ما ولكن نكون عاجزين عن الاجابة عليها في وقت طرح السؤال. والدليل على ذلك نسيان كثير من المعلومات التي كنّا قد تعلمناه عن كثير من الحوادث اما عن طريق المدرسة أو عن أي طريق آخر بعد تعملنا إياه.

ويتأثر احتمال النسيان بمدى أهمية الحدث بالنسبة للشخص وبدرجة تكراره. انه من السهل جداً ان نتذكر بعض الأشياء المهمة كتاريخ التخرج من الثانوية العامة أو تاريخ التخرّج من الجامعة أو تاريخ الزواج أو أول زيارة للعاصمة عمان لمن هو من خارج العاصمة. ولكن من منّا يتذكر تاريخ ثاني زيارة قام بها للعاصمة عمان. وكذلك فانه من السهل جداً لأن نتذكر الأشياء التي تتكرر بحياتنا بشكل مستمر. مثال : العمل أو النشاط الذي نمارسه بشكل مستمر.

ب ـ الاستعداد للاستجابة بدقة Willingness to Respond Accuratelly

ان الشيء الآخر الذي يجب على الباحث اخذه بعين الاعتبار هو تحديد مدى قابلية أو استعداد الشخص المستجوب للاجابة على الأسئلة على فرض انه يمتلك المقدرة في الاجابة على الأسئلة بدقة. ويمكن ان تنعكس عدم الرغبة أو الاستعداد في الاجابة على الأسئلة بما يلي :

(١) رفض المستجوب بالاجابة على سؤال أو مجموعة من الأسئلة. وهذا ينتج عنه خطأ عدم الاستجابة لسؤال أو مجموعة من الأسئلة.

(٢) الاجابة الخاطئة المقصودة وهذا ينتج عنه خطأ قياس.

يوجد هناك ثلاثة أسباب لعدم الاستعداد للاستجابة الدقيقة هي :

(١) قد يكون الظرف غير مناسب للافصاح عن البيانات.

(٢) ان الافصاح عن البيانات قد يكون محرج.

(٣) يعتبر الافصاح عن البيانات خطر حقيقي على سمعة الشخص المستجوب.

يجب أن نتذكر ان حافز الاستجابة على الأسئلة بدقة محدود بالنسبة للشخص المستجوب. ان وجود الشخص المقابل سواء كان في المقابلة الشخصية أو الهاتفية ممكن أن يؤدي الى زيادة الاهتمام من قبل الشخص المستجوب خاصة في الأسئلة المحرجة أو التي تعتبر خطراً عليه.

كما تعتمد درجة استعداد المستجوب في الاجابة على الأسئلة على مدى قناعته بضرورة الحاجة لهذه البيانات للأغراض المشروعة. ان تجميع البيانات التصنيفية يمكن ان يكون مشكلة حقيقية بهذا الشأن. وعليه فان الشخص المستجوب قد يتردد في تزويد بيانات دقيقة اذا ما فوجيء بأسئلة شخصية مثل العمر أو الوظيفة أو الدخل.

٣ ـ تحديد شكل الاستجابة Decide on Response Format

ان اهتمام الباحث يتوجه الى تحديد شكل الاستجابة بعد تحليل المشكلة المتعلقة بمحتوى الأسئلة. ويوجد هناك ثلاثة اشكال للاستجابة على الأسئلة هي :

(١) الأسئلة المفتوحة المنتهية.

(٢) أسئلة الاختيارمن متعدد.

(٣) الأسئلة ذات البديلين فقط.

١ ـ الأسئلة المفتوحة المنتهية Open-Ended Questions

تتطلب الأسئلة المفتوحة المنتهية من المستجوبين تزويد إجابتهم على السؤال. وغالباً ما يطلق على هذا النوع من الأسئلة بأسئلة الاجابة الحرة. وعليه فيمكن طرح السؤال التالي على المستجوبين : « ما هي رد فعلك بالنسبة لكثرة حوادث السير في الفترة الأخيرة؟ » فاذا كانت المقابلة بواسطة البريد، فانه يتم ترك فراغ محدد لكتابة الاجابة على السؤال. أما اذا كانت المقابلة شخصية أو هاتفية فانه يمكن للشخص المستجوب من الاجابة على السؤال شفهياً للشخص الباحث بحيث يقوم الشخص الأخير بكتابة الاجابة على السؤال لاحقاً.

فوائد الأسئلة المفتوحة المنتهية.

أ ـ تساعد على التعبير عن الاتجاهات العامة.

ب ـ تساعد في بناء جسر من التعاون ما بين الشخص الباحث والمستجوب من خلال الاجابة على بعض الأسئلة الدقيقة.

جـ ـ تأثير الأسئلة المفتوحة المنتهية على الاستجابة أقل من الأشكال الأخرى.

د ـ يمكن أن تزود الشخص الباحث بنظرة اعمق أو توضيح أشمل عن بعض المتغيرات.

- سلبيات الأسئلة المفتوحة المنتهية.

أ - ان السلبية الرئيسية للأسئلة المفتوحة المنتهية هي وجود الفرصة الفعلية لتحيز الشخص الباحث. ويمكن أن تنتج فرصة التحيز عن مصدرين هما :

١ - اذا كان الشخص المقابل لا يكتب اجابات الشخص المستجوب مباشرة، فانه يمكن أن يلغي جميع جوانب المقابلة التي لا تبدو ضرورية عند كتابتها.

٢ - اذا كان المقابل بطيء الكتابة أو لا يأخذ رؤوس أقلام عن اجابات الشخص المستجوب، فانه سوف يفشل في كتابة الاجابات بسبب قيد الوقت المفروض عليه.

وكلما زادت الحاجة لتلخيص البيانات وتنقيحها من قبل الشخص الباحث، كلما زاد الفرق ما بين الاجابات المسجلة والاجابات الحقيقية. وللتخلص من مشكلة التحيز الناتجة عن تسجيل اجابات الأشخاص المستجوبين في حالة استخدام وسيلة المقابلة الشخصية أو الهاتفية، فانه يمكن للشخص الباحث من استخدام أداة تسجيل لتسجيل البيانات مباشرة من الشخص المستجوب للتقليل من أخطاء اما القياس أو عدم الاستجابة.

ب - ان السلبية الرئيسية الأخرى لهذا النوع من الأسئلة ناتجة عن الوقت والتكاليف المترتبة على عملية ترميز الاجابات. فاذا كانت الدراسة كبيرة فسوف تكون هناك حاجة كبيرة لاجراءات الترميز والترقيم لتلخيص البيانات المتناثرة ووضعها بشكل معين يمكن من خلاله القيام بتحليل البيانات ومن ثم عرضها. ويمكن أن تشكل تكاليف ووقت عملية الترميز جزء كبير من التكاليف الكلية لمشروع البحث.

جـ ان السلبية الأخرى لهذا النوع من الأسئلة هي اعطاء وزن أكثر من اللازم وبطريقة ضمنية للأشخاص الذين يفصلّون في الاجابات بشكل كبير أو الأشخاص الذين يطرحون تساؤلات كثيرة من خلال اجاباتهم.

د ـ لا يعتبر هذا النوع من الأسئلة مناسباً في حالة استخدام الطريقة البريدية لأن الأشخاص المستجوبين غالباً ما يقوموا بكتابة الاجابات بشكل مختصر مقارنة مع ما يتكلموا به. هذا بالاضافة الى صعوبة قراءة بعض كتاباتهم في بعض الأحيان.

وبشكل عام، تعتبر طريقة الأسئلة المفتوحة المنتهية أكثر ملائمة بالنسبة للأبحاث الاستكشافية وتصميم البحث بهدف تطوير أسئلة موجهة لحل مشكلة البحث بطريقة افضل.

ب ـ أسئلة الاختيار من متعدد Multiple-Choice Questions

تتطلب اسئلة الاختيار من متعدد من الشخص المستجوب اختيار الاجابة من الاجابات المتعددة التي تتبع كل سؤال من اسئلة الاستبيان. ويمكن أن يطلب الشخص الباحث من الشخص المستجوب اختيار اجابة أو أكثر من الاجابات المرفقة.

فوائد أسئلة الاختيار من متعدد.

١ ـ امكانية هذا النوع من الأسئلة في الغاء كثير من السلبيات المتعلقة بطريقة الأسئلة المفتوحة المنتهية. واهمها هو تقليل التحيز الذي ينتج عن طريق الشخص المقابل وتقليل التكلفة والوقت اللازم لاعداد البيانات.

٢ ـ سهولة التعامل مع هذا النوع من الأسئلة لغايات تحليل البيانات وعرض النتائج.

٣ ـ صعوبة الحصول على تعاون الشخص المستجوب في بعض الأحيان الا عن طريق استخدام أسئلة الاختيار من متعدد.

سلبيات أسئلة الاختيار من متعدد.

١ ـ يحتاج تصميم أسئلة الاختيار من متعدد الى وقت كثير وتكلفة عالية نسبياً مقارنة مع تصميم الأسئلة المفتوحة المنتهية.

٢ ـ يمكن ان تؤدي أسئلة الاختيار من متعدد الى التحيز الناتج عن ترتيب الاجابات المزودة في الأسئلة.

جـ ـ الأسئلة ذو الاختيارين فقط

تعتبر الأسئلة ذو الطرفين حالة خاصة من أسئلة الاختيار من متعدد والتي تسمح للشخص المستجوب اختيار اجابة من بديلين فقط، مثل نعم أو لا، أوافق أو اعارض، وما الى ذلك. وغالباً ما يكون هناك اختيار محايد بالاضافة الى البديلين الموجودين، مثل « لا رأي» أو لا أدري» .

فوائد الأسئلة ذو الاختيارين

ان ايجابيات هذا النوع من الأسئلة هي نفس ايجابيات أسئلة الاختيار من متعدد والتي تشمل ما يلي :

(١) سرعة الاجابة.

(٢) سهولة تدبير الأسئلة.

(٣) تقليل فرصة تحيّز الشخص الباحث.

(٤) سهولة ترميز وترقيم الاجابات واعدادها وتحليلها.

سلبيات الأسئلة ذو الاختيارين.

١ ـ صعوبة تعبير الشخص المستجوب عن شعوره أو الشيء الذي يحس به تجاه السؤال عن طريق إجابتين فقط.

٢ ـ إجبار الشخص المستجوب للاختيار اجابة من اجابتين فقط ممكن أن يؤدي الى خطأ في القياس.

٣ ـ ان الاجابة السلبية أو الايجابية عن السؤال يمكن ان يكون لها تاثير قوي على طبيعة الاجابات.

٤ ـ تحديد كلمات السؤال Decide on Question Wording

يتكون قلب الاستبيان من الأسئلة التي يتم طرحها من خلاله. ان هذه الأسئلة تمثل العلاقة ما بين البيانات والمعلومات المطلوبة للدراسة. لذلك فانه من المهم جداً ان يحمل السؤال نفس المعنى لكل من الشخص الباحث والشخص المستجوب لكي لا يظهر خطأ القياس في نتائج دراسة البحث.

ويمكن اعتبار الخطوط الاسترشادية التالية عند تصميم كلمات الأسئلة :

(أ) استخدام كلمات بسيطة.

(ب) استخدام كلمات واضحة.

(جـ) تجنب الأسئلة القيادية.

(د) تجنب الأسئلة المتحيزة.

(هـ) تجنب التقدير.

(و) تجنب الأسئلة المزدوجة.

أ ـ استخدام الكلمات البسيطة Use Simple Words

يجب ان تكون الكلمات المستخدمة في الاستبيان منسجمة مع مستوى المصطلحات بالنسبة للمستجوبين. فالأسئلة المصممة للأطفال يجب ان تحتوي على كلمات أبسط مقارنة مع كلمات أسئلة الاستبيان الموجهة لحملة الماجستير. اما اذا كان مجتمع الدراسة يحتوي على فئات مختلفة من الناس، فانه يجب كتابة أسئلة الكلمات باستخدام كلمات يمكن فهمها من اقل الناس معرفة.

ب ـ استخدام الكلمات الواضحة Use Clear Words

ان الكلمات الواضحة هي الكلمات التي لا تحمل إلا معنى واحد ومعروف لكل عناصر مجتمع الدراسة. ولكن لسوء الحظ، ان عملية اختيار كلمات واضحة أو غير غامضة قد تكون اصعب مما هو متوقع. ان هناك العديد من الكلمات التي تظهر وكأنها واضحة لكل فرد قد تحمل معنى مختلف لبعض عناصر مجتمع الدراسة أو في مناطق جغرافية مختلفة.

وقد وجد الباحثون بالخبرة ان استخدام بعض الكلمات مثل غالباً أو بانتظام أو عادة أو باستمرار جميعها كلمات غير واضحة أو غامضة. لذلك يتوجب على الباحث وقبل اختيار الكلمة ان يقوم بطرح الأسئلة التالية :

أ ـ هل الكلمة المستخدمة تعطي المعنى المقصود ؟

ب ـ هل للكلمة المستخدمة أي معنى آخر غير المعنى المقصود ؟

جـ ـ اذا كان الجواب على السؤال رقم (٢) بنعم، فهل سياق الكلام يجعل المعنى المقصود واضح ؟

د ـ هل هناك كلمات اخرى بديلة مقترحة ؟

جـ ـ تجنب الأسئلة القيادية Avoid Leading Questions

السؤال القيادي هو ذلك السؤال الذي يترك اشارة معينة للشخص المستجوب بحيث تحدد نوع الاجابة بالنسبة للسؤال. وغالباً ما تعكس الأسئلة القيادية وجهة نظر الشخص الباحث بالنسبة لاجابة السؤال أو مجموعة الأسئلة. وغالباً ما تؤدي الأسئلة القيادية الى وقوع خطأ القياس وبشكل مستمر في نتائج البحث.

ان طرح السؤال التالي : « هل تمتلك جهاز تلفزيون ماركة توشيبا » قد يكون سؤالاً قيادياً مقارنة مع السؤال التالي : « ما هو نوع جهاز التلفزيون الذي تمتلكه؟ » ان

استخدام اسم الشركة في السؤال قد يسبب في إيهام المستجوب بأن هذه الشركة هي الشركة التي تدعم دراسة البحث. وفي هذه الحالة فانه قد يتولد نزعة عند الشخص الباحث في اظهار شعور ايجابي تجاه الشركة التي تدعم مشروع البحث، وهذا بدوره يؤدي الى الوقوع في خطأ القياس.

د ـ تجنب الأسئلة المتحيزة Avoid Biasing Questions

تتضمن الأسئلة المتحيزة كلمات أو عبارات عاطفية بحيث تمتلك اقتراح شعوري إما بالقبول أو بالرفض . ان معظم الباحثين يدركون التأثير التحيزي للأسئلة التي تبدأ بـ « هل توافق على الاقتراح المقدم من .... » أو « هل تؤمن بما يجب عمله بالنسبة لشركات احتكار القلّة بشأن .... ». ولسوء الحظ، ان تأثير استخدام الكلمات أو العبارات المتحيزة في الاستبيان يكون أكثر خداعاً مما تقترحه هذه الأمثلة. ان استخدام الكلمات أو العبارات المتخيزة في أسئلة الاستبيان قد يولد نزعة عاطفية اما ايجابية أو سلبية عند الشخص المستجوب، وهذا بدوره يؤدي الى الوقوع في خطأ القياس بالنهاية.

هـ ـ تجنب التقدير Avoid Estimation

يجب ان تصمم الأسئلة بطريقة يكون فيها الشخص المستجوب قادراً على اعطاء اجابة واضحة وليس عن طريق التقدير أو عمل التعاميم. ومن الأمثلة على هذا النوع من الأسئلة : ما هو عدد فناجين القهوة التي تقوم بشربها بالسنة؟ ان هذا السؤال يتطلب من الشخص المستجوب تحديد عدد فناجين القهوة التي يشربها باليوم الواحد ومن ثم يقوم بضرب ذلك بعدد أيام السنة. ان نتيجة البحث تكون أكثر دقة فيما لو كان السؤال كما يلي : كم فنجاناً من القهوة تشرب باليوم الواحد؟ وعليه فانه يتم معرفة عدد الفناجين التي يقوم بشربها عن طريق الضرب بعدد أيام السنة.

و ـ تجنب الأسئلة المزدوجة Avoid Double-Barrelled Questions

يحدث هذا النوع من الأسئلة اذا كان المطلوب من الكلمات المستخدمة في السؤال اجابتين وليس اجابة واحدة. ومن الأمثلة على هذا النوع من الأسئلة ما يلي : «ما هو تقديرك لسرعة السيارة وتسارعها» اننا نلاحظ هنا ان هذا السؤال يتكون من شقين هما السرعة والتسارع. وكقاعدة ان أي سؤال يتضمن حرف العطف (و) يجب مراجعته للتأكد من عدد الاجابات المطلوبة.

٥ ـ تحديد تسلسل الأسئلة Decide on Question Sequence

ان الخطوة التي تتبع تحديد كلمات أسئلة الاستبيان هو تحديد تسلسل الأسئلة في الاستبيان. ان تسلسل الأسئلة يمكن أن يؤثر على طبيعة اجابات الأشخاص المستجوبين ويكون هذا هو السبب في وقوع أخطاء خطيرة في نتائج البحث. ويوجد هناك مجموعة من الخطوط الاسترشادية التي يمكن استخدامها من قبل الباحثين المبتدئين عند تصميم الاستبيان. وهذه الخطوط الاسترشادية هي :

أ ـ استخدام الأسئلة السهلة والممتعة كأسئلة افتتاحية

يجب على الأسئلة الافتتاحية أن تستقطب اهتمام واحترام المستجوب مباشرة والا فان المستجوب قد ينهي المقابلة. وغالباً ما تتعلق الأسئلة الافتتاحية بالمعلومات المطلوبة، بحيث يكون الغرض الوحيد منها هو كسب تعاون المستجوب وتحقيق الصلة معه. وعليه يمكن ان يبدأ الاستبيان بالأسئلة السهلة التي تطلب من المستجوب التعبير عن اتجاهاته، لأن معظم الناس يرغبون في التعبير عن شعورهم ويقومون بذلك بسهولة. ان اتباع هذا المبدأ يمنح الشخص المستجوب الثقة على أنه قادر على اجابة ما يتبع من أسئلة.

ب ـ البدء بالأسئلة العامة أولاً

يجب كتابة الأسئلة العامة في اي دراسة قبل الأسئلة الخاصة. اعتبر الآن السؤالين التاليين : « ما هي الاعتبارات المهمة التي تأخذها بعين الاعتبار عند شرائك سيارة؟ » و « عند شرائك سيارة، هل موضوع اللون مهم لك؟ » اذا تم طرح الأسئلة بطريقة عكسية فان موضوع اللون قد يظهر عند الاجابة عن السؤال الأول بصورة أكبر مقارنة مع الترتيب الحالي الذي ظهرت عليه الأسئلة هنا. وكنتيجة ان طرح الأسئلة العامة أولاً والخاصة تقلل من فرصة التحيز الناتج عن التسلسل.

جـ ـ وضع الأسئلة الصعبة وغير الممتعة بالنهاية

انه من الصعب جداً أن يقوم الشخص المستجوب برفض الأسئلة الشخصية والصعبة بعد بناء نوع من الصلة ما بينه وبين الشخص الباحث وعملية البحث. لذلك يفضل دائماً ترتيب الأسئلة الصعبة والحساسة والمحرجة في نهاية الاستبيان.

د ـ ترتيب الأسئلة بترتيب منطقي

يجب ترتيب عملية طرح الأسئلة ترتيباً منطقياً من وجهة نظر المستجوب. ان الهدف من تصميم تسلسل الأسئلة هو لتسهيل عملية معالجة البيانات وللتأثير على تعاون الشخص المستجوب وبناء علاقات معه.

٦ ـ تحديد الخصائص المادية.

ان المظهر المادي الخارجي للاستبيان قد يكون مؤثراً في ضمان التعاون مع الشخص المستجوب خاصة في حالة استخدام المقابلة البريدية. ان نوعية ورق الاستبيان ونوع الطباعة غالباً ما يحدد ردة فعل المستجوب الأولى نحو الاستبيان. كما انه من المهم جداً وضع اسم المؤسسة التي تدعم البحث واسم الدراسة بوضوح في بداية الاستبيان.

٧ ـ تنفيذ الاختبار الأولي والمراجعة والشكل النهائي.

يجب على الباحث ان يقوم بفحص الاستبيان فحصاً أولياً قبل ان يوزعه على مجتمع الدراسة بشكله النهائي. ويشير الفحص الأولي الى فحص جانب أو أكثر من جوانب تصميم البحث. ان الاهتمام هنا يكون في محض تصميم الاستبيان للبحث في امكانية تحسينه وبغض النظر عن مهارة الشخص الباحث. ان معظم الاستبيانات بحاجة الى امتحان أولي ومراجعة واحدة على الأقل قبل تنفيذ الدراسة الميدانية.

ان عدد الأشخاص الذين يتم مقابلتهم لاجراء الاختبار الأولي للاستبيان يكون غالباً ما بين ١٥ ـ ٣٠ فرداً. ويجب أن تكون هذه العينة مشابهة من حيث الخصائص لعينة الدراسة الرئيسية.

وكلما كان التغيير المطلوب على اسئلة الاستبيان ذو معنى، كلما كان هناك حاجة لعمل اختبار أولي آخر. اما اذا كان الاختبار الأولي يقترح اجراء بعض التعديلات الخفيفة، فان الاستبيان يكون جاهز بشكله النهائي لاجراء الدراسة الميدانية.

ـ أشكال الملاحظة.

ان تصميم أشكال الملاحظة أسهل من الاستبيانات لان عملية طرح الأسئلة تكون محذوفة. الا ان هناك بعض القضايا المهمة في بناء أشكال الملاحظة. يجب على الشخص الباحث ان يحدد وبوضوح انواع الملاحظات التي سيقوم بها وكيفية قياسها. ويمكن ان تتضمن عملية القياس وجود شخص ملاحظ مع أشكال الملاحظة او وسيلة تسجيل آلية او كلاهما.

يجب ان يبدأ تصميم أشكال الملاحظة من قائمة المعلومات المطلوبة والتي يجب ان تحدد وبوضوح جميع الجوانب السلوكية المراد مراقبتها. وغالباً ما يكون من المفيد تصنيف المعلومات السلوكية المطلوب ملاحظتها كما يلي:

١ ـ من هو المراد ملاحظته؟ ـ البائع، المشتري، المستهلك.

٢ ـ ما هو الشيء الذي نريد ملاحظته؟ نوع السلعة، حجمها، السعر، تاثير سلوك معين على الأطفال او الرجال.

٣ ـ متى نقوم بالملاحظة؟ ساعات الشراء، وقت الاستراحة.

٤ ـ أين يجب ان تتم الملاحظة؟ بالمنزل او بالمحل التجاري.

يجب ان تكون اشكال الملاحظة سهلة الاستخدام. ويجب تصميم اشكال الملاحظة بحيث تكون قادرة على ملاحظة السلوك. كما يجب ان يسمح التصميم للشخص الباحث من تسجيل السلوك بالتفصيل. وكما يجب ان تتبع الخصائص المادية لشكل الملاحظة نفس القوانين المستخدمة في الاستبيان. واخيراً تحتاج اشكال المراقبة نفس الدرجة من الاختبار الأولي والمراجعة كما هو الحال بالنسبة للاستبيان.

أسئلة

س١ : ما هو الدور الذي يلعبه الاستبيان في مشروع البحث ؟

س٢ : ما هي المكونات الرئيسية للاستبيان ؟

س٣ : ما هي القراءات التي يجب أن تسبق مرحلة تصميم الاستبيان ؟

س٤ : لماذا يكون بعض المستجوبين غير مستعدين للاجابة على بعض الأسئلة بدقة ؟

س٥ : أذكر ايجابيات وسلبيات طريقة الأسئلة المفتوحة المنتهية ؟

س٦ : أذكر ايجابيات وسلبيات طريقة الاختيار من متعدد ؟

س٧ : أذكر الخطوط الاسترشادية التي يجب استخدامها عند تصميم كلمات الأسئلة ؟

الفصل التاسع

اعداد البيانات

Data Processing

اعداد البيانات
Data Processing

ان اهتمام الباحث بعد الحصول على البيانات من خلال طرق جمع البيانات المختلفة يتحول الى تحويل هذه البيانات الى شكل يمكّن الشخص الباحث من تحليل البيانات التي تحتوي عليها أداة جمع البيانات. لذلك يتوجب على الشخص الباحث معرفة ما يجب عمله لتحقيق هذه المهمة.

يطلق على الوظائف الضرورية لاعداد أشكال جمع البيانات الأولية بالأدوات لغايات اعداد هذا الفصل.

ـ بعض المبادىء الأساسية

ان المهمة الأساسية من اعداد البيانات هو تحويل البيانات الأولية في أداة جمع البيانات الى شكل قابل للقراءة بالنسبة لجهاز الحاسوب. وبعد ذلك نستطيع استخدام اجراءات تحليل البيانات لاستخراج معلومات من البيانات. ولكن قبل أن نبين كيفية عمل ذلك، فان هناك بعض المبادىء التي نحن بحاجة الى تفهمها.

أ ـ الحالة Case

تشير الحالة الى وحدة التحليل المخصصة بالنسبة للدراسة. وغالباً ما تشير وحدة التحليل الى الشخص المستجوب. لذلك يعتبر كل فرد من الأفراد المستجوبين كحالة، ومجموع عدد الحالات يساوي عدد أفراد العينة.

ب ـ بطاقة الحاسوب The Computer Card

يجب على الشخص الباحث القيام بتحويل البيانات المزودة من أداة البحث الى شكل مقروء لجهاز الحاسوب، وتعتبر بطاقة الحاسوب أفضل وسيلة يمكن استخدامها لهذا الغرض.

جـ ـ صندوق البيانات Data Deck

يتكون صندوق الحاسوب من جميع بطاقات الحاسوب الضرورية لتمثيل البيانات الموجودة في الستبيانات. مثال : اذا كان الاستبيان الواحد يحتاج الى بطاقتي حاسوب لتمثيل البيانات الموجودة فيه وكان حجم عينة الدراسة هو ٥٠٠ شخص، فاننا سوف نحتاج الى ٢ X ٥٠٠ = ١٠٠٠ بطاقة لصندوق البيانات. ان البيانات الموجودة في صندوق البيانات هي التي يتم ادخالها الى جهاز الحاسوب.

د ـ مخزن البيانات Data Storage

بعد ادخال البيانات الى جهاز الحاسوب، فان الباحث سوف يقوم بتخزين البيانات الموجودة في صندوق البيانات اما في ذاكرة الحاسوب أو على شريط الحاسوب. وبعد تخزين البيانات فلا يوجد هناك أي مشكلة فيما لو فقد الباحث صندوق البيانات. كما ان استخدام البيانات المخزنة بالنسبة لجهاز الحاسوب في عمليات التحليل اللاحقة يكون أسهل.

هـ ـ مصفوفة البيانات Data Matrix

يمكن التفكير في البيانات المخزونة اما في ذاكرة الحاسوب أو على شريط الحاسوب كمصفوفة بيانات (انظر الشكل ٩ ـ ١). فكل صف في المصفوفة يمثل حالة، وكل عمود يمثل متغير، لاحظ ان عدد الصفوف بالمصفوفة يساوي عدد الحالات (ن) كما ان

عدد الأعمدة يساوي عدد المتغيرات (م). لذلك حجم مصفوفة البيانات هو ن X م. ان البيانات الممثلة في مصفوفة العينة هي قيم ذات علاقة بثلاثة متغيرات لمجموعة محددة من الحالات، وهذه المتغيرات هي المستوى الجامعي وعلامة أساليب البحث العلمي وعدد أفراد الأسرة. لاحظ ان كل متغير ومهما كان عدد الخانات التي يتطلبها ياخذ عمود واحد فقط من مصفوفة البيانات. ان عدد الحالات والمتغيرات في مصفوفة البيانات يتحدد بسعة الحاسوب. انه من الطبيعي في الحياة العملية أن نشاهد مصفوفة بيانات ذات حجم ٢٠٠٠ X ٢٠٠ أو أكثر. ان البيانات التي يستخدمها جهاز الحاسوب في التحليل هي بيانات مصفوفة البيانات.

وبعد التعريف بهذه المصطلحات الأساسية، فاننا سوف نقوم الآن بدراسة كل خطوة من خطوات اعداد البيانات.

ـ خطوات اعداد البيانات Data Processing Flow

يتطلب اعداد البيانات اتباع الخطوات التالية :

(١) تحديد أداة جمع البيانات.

(٢) التدقيق.

(٣) الترميز.

(٤) تحويل صندوق البيانات الى شكل مقروء بالنسبة لجهاز الحاسوب.

(٥) تخزين صندوق البيانات اما في ذاكرة الحاسوب أو على شريط حاسوب.

شكل (٩ ـ ١)

مبدأ مصفوفة البيانات

			المتغيرات من ١ ← م			
			عدد الأعمدة			
			١	٢	....	م
الحالات		١	٢	٨٠		٢
من ١ ← N	عدد الصفوف	٢	٤	٧٥		٤
		٣	٢	٦٠		٣
		٤	٣	٩٠		٨
		.				
		.				
		.				
		ن	١	٥٠		٥

المتغير رقم ١ = المستوى الجامعي (١ = سنة أولى ، ... ، ع = سنة رابعة) .

المتغير رقم ٢ = علامة أساليب البحث (من ١٠٠) .

المتغير رقم ٣ = عدد أفراد الأسرة .

١ ـ تحديد أداة جمع البيانات Identification of Acceptable Instruments

يتوجب على الشخص الباحث تقييم جميع أودات جمع البيانات ومن ثم اختيار أكثر الأدوات قبولاً للاستخدام في الدراسة. وتختلف معايير مقارنة ادوات جمع البيانات من دراسة الى أخرى. ان أداة جمع البيانات الأمثل للشخص الباحث هي تلك

الأداة التي نستطيع من خلالها تعديل اي خطأ في اي سؤال من أسئلة الاستبيان ضمن وقت معقول وتكلفة مقعولة.

٢ ـ التدقيق Editing

ان المقصود بالتدقيق هو مراجعة أدوات جمع البيانات لضمان أعلى دقة ووضوح. لذلك فانه من المهم القيام بعملية التدقيق بشكل مستمر. اذا كانت الدراسة صغيرة فانه يمكن للشخص الباحث القيام بوظيفة التدقيق بنفسه. أما اذا كانت الدراسة كبيرة فان وظيفة التدقيق قد تحتاج الى عدة مدققين. وعند القيام بعملية التدقيق، يجب الاهتمام بما يلي :

أ ـ الوضوح Clearness

يجب ان تكون البيانات واضحة حتى يتمكن الباحث من ترميزها بشكل مناسب لاحقاً. في بعض الأحيان يمكن تصحيح بعض الاجابات غير الواضحة عن طريق الاتصال مع الشخص لذي قام بتسجيل الاجابة، اما في بعض الأحيان الأخرى فيمكن تصحيح الاجابة بناء على الأجزاء الأخرى من أداة جمع البيانات. فاذا لم يتوفر اجابة واضحة ودقيقة لأحد الأسئلة فيجب الاشارة الى الاجابة كبيانات ناقصة. بمعنى آخر، يجب على المدقق أن يزيل الغموض من البيانات المسجلة حتى يتمكن الشخص الذي سوف يقوم بعملية الترميز من معرفة ما سوف يقوم به بالتحديد

ب ـ الكمال في الاجابة Completeness

يمكن معالجة الأسئلة التي لا تحتوي على اجابة بإحدى الطرق التالية :

(١) يمكن للشخص المدقق الاتصال مع الشخص المقابل أو الباحث لتحديد ما اذا تم نسيان الاجابة على السؤال من قبل الشخص المستجوب أو ان الشخص المقابل قد غفل عن تدوين الاجابة. ان الخطورة في عمل ذلك هو عدم مقدرة الشخص المقابل في تذكر المقابلة المعنية بالتحديد.

(٢) تصنيف تلك البيانات كبيانات مفقودة.

(٣) اذا تبين للمدقق فقدان كثير من البيانات فيجب اعادة جمع البيانات مرة أخرى بنفس الأداة او استبدالها بأداة أخرى.

جـ ـ الدقة Accuracy

يجب على الشخص المدقق ان يكون يقظاً لأي دليل من أدلة عدم دقة البيانات. وأهم شيء يجب ان يتنبه له المدقق هنا هو احتمال تحيز أو غش الشخص المقابل. ويمكن للمدقق من ملاحظة ذلك النشاط من خلال نمط الاجابات في أداة جمع البيانات لبعض المقابلات.

د ـ توضيح الاجابات Response Clarification

تكون الاجابات للأسئلة المفتوحة المغلقة صعبة التفسير بوضوح بالنسبة للمدقق اما بسبب استخدام الاختصارات بشكل كبير أو وجود بعض الكلمات الغامضة. فإما ان يعين المدقق معنى لتلك الاجابة أو يقوم بسؤال الشخص المقابل عن معنى تلك الاجابة. ولكن احتمال خطر الوقوع بالخطأ عالي في كلا الحالتين. ولكن الدراسة الميدانية الجيدة تمنع من ظهور كثير من المشاكل.

كما ان بعض الصعوبات تتحقق عند الاجابة على الأسئلة بطريقة تختلف عن تعليمات أداة جمع البيانات. وغالباً ما تحدث هذه الأخطاء بطريقة المقابلة البريدية. فعلى سبيل المثال : اذا تم الطلب من الشخص المستجوب اختيار رقم واحد على ميزان مقياس خماسي وقام باختيار رقمين هما ٣ و ٤، فأي الاجابة هي المقصودة بالنسبة للشخص المستجوب، وهل يقصد الشخص المستجوب اختيار الاجابة ٣٫٥. لهذا يجب أن يقرّر الباحث اما ان يختار أحد الاجابتين ٣ أو ٤ أو يقوم بتسجيل تلك النوع من البيانات كبيانات مفقودة.

٣ ـ الترميز Coding

يتضمن الترميز تعيين رقم لكل عمود من أعمدة بطاقة الحاسوب لتمثيل الاجابات في أداة جمع البيانات. وغالباً ما تتم عملية الترميز في طريقة المقابلة البريدية قبل تنفيذ العمل الميداني. فعلى سبيل المثال يمكن ترميز اجابات السؤال الذي يهتم بالجنس كما يلي :

الجنس ١ ـ ذكر.

٢ ـ أنثى.

ان الأرقام الموجودة على جهة اليمين تشير الى موضوع الترميز. فالرقم (١) يعني ان جنس المستجوب هو ذكر والرقم (٢) يعني ان جنس المستجوب هو أنثى. ان عملية الترميز بالنسبة للأسئلة المفتوحة المغلقة سواء كانت في المقابلة البريدية أو الهاتفية أو الشخصية تكون أكثر صعوبة مقارنة مع الأسئلة ذات الاختيار من متعدد. ان الشخص المقابل هو الشخص الذي يقوم بتسجيل الاجابات الحرفية لمجموعة الأشخاص المستجوبين. فكيف يتم تحويل مثل هذا النوع من البيانات الى شكل مقروء. يوجد هناك طريقتين لحل هذه المشكلة. تتمثل الطريقة الأولى في اعداد طريقة مناسبة للترميز قبل الانتهاء من الدراسة الميدانية. وحتى يتمكن الشخص الباحث من القيام بذلك فانه لا بد من وجود نتائج دراسات سابقة بحيث تعمل كدليل عند القيام بعملية الترميز. لذلك فان المهمة الرئيسية للشخص الباحث هي تدريب مجموعة الأشخاص الذين يقومون بعملية الترميز حتى يتمكنوا من تحويل الاجابات الشفهية الى شكل مقروء بالنسبة لجهاز الحاسوب.

أما بالنسبة للطريقة الثانية فهي انتظار حتى عودة أدوات جمع البيانات من الدراسة الميدانية قبل اعداد خطة الترقيم أو الترميز. وهنا يقوم الباحث بتدوين (٥٠) الى (١٠٠) اجابة لكل سؤال. وبعد ذلك يقوم الشخص الباحث بفحص هذه الاجابات واتخاذ القرار المناسب بالنسبة للتصنيف المناسب لتلخيص البيانات.

٤ ـ تحويل البيانات الى شكل مقروء Convert Data Deck into a Computer Readable Form.

بعد اجراء عملية الترقيم يجب عرض صندوق البيانات الى جهاز الحاسوب بطريقة يمكن استخدامها من قبل برامج التحليل الحاسوبية. ان معظم البرامج التي يحتاجها الشخص الباحث غالباً ما تكون على شكل رزم احصائية. ان اسم إحدى الرزم الاحصائية هو الرزم الاحصائية للعلوم الاجتماعية (SPSS) .

٥ ـ التخزين Storing

بعد اعداد مجموعة من البيانات بصورة كاملة الى شكل مقروء يجب تخزين صندوق البيانات اما في ذاكرة الحاسوب الرئيسية أو في شريط حاسوب. ويفضل حفظ البيانات على شريط حاسوب آخر يحفظ في مكان امين كنوع من الاحتياط. وبعد ذلك تكون البيانات جاهزة للتحليل.

أسئلة

س١ : كيف يتم تمثيل أسئلة الاستبيانات في بطاقة الحاسوب ؟

س٢ : كيف يتم تحديد أدوات جمع البيانات غير المناسبة لعملية اعداد البيانات ؟

س٣ : ماذا يجب على المدقق عمله عند فحص أداة جمع البيانات ؟

س٤ : كيف يتم معالجة الاجابات المتعددة ؟

الفصل العاشر

تحليل البيانات

Data Analysis

تحليل البيانات
Data Analysis

ـ مقدمة

بعد تحويل البيانات التي تم تجميعها باستخدام وسائل جمع البيانات المختلفة الى شكل مقروء بالنسبة لجهاز الحاسوب كما تبين في الفصل السابق، يجب ان يتحول الاهتمام عند الشخص الباحث الى تحليل هذه البيانات لاستخلاص المعلومات المطلوبة لموضوع القرار. ان التحليل غير المناسب للبيانات يمكن ان يكون مصدراً خطيراً للأخطاء غير العينية. لذلك ان الهدف الرئيسي من كتابة هذا الفصل هو لمحاولة تحديد الاستخدام الأمثل لكل اختبار تحليلي حتى نتمكن استخلاص المعلومات من البيانات بطريقة صحيحة. وحتى نتمكن من ذلك فانه لا بد من تحديد ظروف استخدام كل اختبار أو كل طريقة مع اعطاء بعض الأمثلة.

يعتقد بعض الباحثين بأن تحليل البيانات هو أهم جانب في البحث الاداري. ولكن المنطق الذي سوف نعتبره هنا هو ان تحليل البيانات الاحصائي المعقد لا يعوض الضعف في تحديد المشكلة أو تصميم الدراسة السيء أو اختيار العينة الخاطيء أو الضعف بالقياس أو اعداد البيانات السيء. ان تحليل البيانات هو أحد النشاطات العديدة التي تحتاجها عملية البحث والتي يجب أن تنفذ بشكل صحيح للحصول على معلومات مناسبة لاتخاذ القرار. لذلك يجب أن تتم عملية التحليل بشكل مناسب أيضاً للتمكن من الحصول على المعلومات المطلوبة لحالة القرار.

ويجب أن نبدأ هذا الفصل بالتمييز ما بين اجراءات تحليل البيانات أحادية وثنائية ومتعددة المتغيرات. وبعد ذلك ننتقل الى التمييز ما بين الاجراءات التي يكون الهدف

من استخدامها هو وصف البيانات والتي يطلق عليها بالاجراءات الاحصائية الوصفية والاجراءات التي يكون الهدف من استخدامها هو استخلاص المعلومات واتخاذ القرارات والتي يطلق عليها بالاجراءات الاحصائية التحليلية أو الاستقرائية. واخيراً، فاننا سوف نقوم بشرح عدداً من الاختبارات الاحصائية الوصفية والتحليلية.

ولكن قبل البدء بشرح الاجراءات الاحصائية الوصفية أو التحليلية، دعنا نقوم بشرح طرق عرض البيانات.

ـ طرق عرض البيانات

ان الطرق المستخدمة في تجميع البيانات في العلوم الادارية تعطي بيانات كثيرة، وهذه البيانات عبارة عن مشاهدات على شكل أرقام، ربما يكون عددها كبير جداً. فاذا تم عرض هذه البيانات بطريقة انشائية، فانه قد يكون من الصعب استيعابها والمقارنة بين مفرداتها من أجل استخلاص النتائج أو المعلومات المطلوبة منها. لذلك فانه من الضروري عرض هذه البيانات بطريقة يمكن من خلالها اتسخلاص المعلومات من البيانات المتوفرة، ومن هذه الطرق :

١ ـ طريقة الجداول

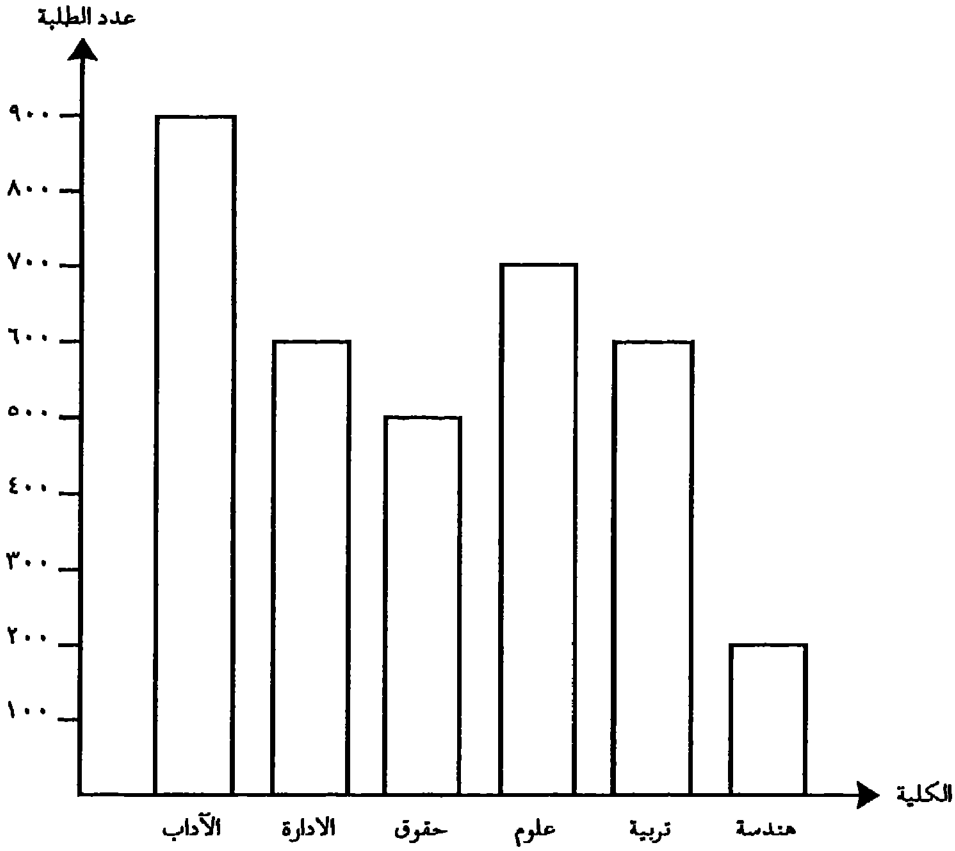
طريقة الجداول عبارة عن وضع البيانات في جدول. وكثير ما تستعمل الجداول في عرض تغيّر ظاهرة مع مرور الزمن أو مع مسميات كالدول أو الشركات أو كلاهما معاً. فاذا كانت البيانات مثلاً عبارة عن اعداد الطلبة المقبولين في جامعة مؤتة في الفترة ما بين ١٩٨٦ ـ ١٩٩٠ حسب الكليات مثلاً. فانه يمكن تمثيل هذه البيانات على شكل جدول كما يلي :

الكلية	عدد الطلبة
الآداب	٩٠٠
الاقتصاد والادارة	٦٠٠
الحقوق	٥٠٠
العلوم	٧٠٠
التربية	٦٠٠
الهندسة	٢٠٠
المجموع	٣٥٠٠ طالب

٢ ـ طريقة الأعمدة أو المستطيلات

يتلخص عمل هذه الطريقة بوضع المسمّيات على محور افقي أو عمودي ومن ثم رسم مستطيل على كل مسمّى بحيث يكون ارتفاع كل مستطيل ممثلاً للقيمة المقابلة لذلك المسمّى وذلك باستعمال مقياس رسم مناسب.

وتستخدم هذه الطريقة لغايات المقارنة بين قيم الظواهر حسب الزمن أو المسمّيات. وكذلك يمكن استخدامها للمقارنة بين قيم الظواهر حسب المسميات على مدى عدّة سنوات. لذلك يمكن عرض البيانات الواردة في طريقة الجداول بطريقة الأعمدة وكما يلي :



٣ ـ طريقة الدائرة

أهم استخدامات هذه الطريقة هو لتقسيم الكل الى أجزائه. فاذا أردنا عرض البيانات الواردة في طريقة الجداول بهذه الطريقة، فاننا نمثل مجموع الطلبة بدوائر كاملة ونمثل الطلبة المقبولين في كل كلية بقطاع من الدائرة بحيث يكون قياس زاويته هو :

$$\text{قياس زاوية القطاع} = \frac{\text{عدد الطلبة المقبولين في الكلية}}{\text{مجموع الطلبة}} \times ٣٦٠^{\circ}$$

وعليه يكون قياس زاوية قطاع كلية الآداب هو

$$\text{الزاوية} = \frac{٩٠٠}{٣٥٠٠} \times ٣٦٠^\circ = ٩٢٫٥٨^\circ$$

ـ مراجعة اجراءات تحليل البيانات

ان السؤال الرئيسي الذي يواجه الباحث عند تحليل البيانات هو : «ما هو التحليل الاحصائي المناسب والذي يجب استخدامه في تحليل البيانات؟» ان الاجابة على هذا السؤال تتطلب معرفة الحالات التي تواجه الشخص الباحث وبدقة. ويوجد هناك ثلاثة انواع من الأسئلة التي يمكن ان تساعد الباحث في تحديد الاختبار المناسب، وهذه الأسئلة هي :

(١) ما هو عدد المتغيرات التي يجب تحليلها في كل مرحلة من مراحل التحليل؟

(٢) هل يريد الباحث وصف البيانات فقط أم استخلاص معلومات من بيانات؟

(٣) ما هو مستوى القياس المستخدم، هل هو اسمي أم ترتيبي أم مقياس فترة ؟

* عدد المتغيرات المراد تحليلها.

ان الجانب الأول والذي يجب توضيحه يعود الى عدد المتغيرات الذي يرغب الشخص الباحث بتحليلها في كل مرحلة من مراحل التحليل. فاذا أراد الشخص الباحث اجراء الاختبار على أساس متغير واحد في كل مرحلة فيسمى هذا بالتحليل أحادي المتغير. أما اذا أراد الباحث اختبار متغيرين في كل مرحلة من مراحل التحليل فان هذا يسمى بالتحليل ثنائي المتغير. أما اذا أراد اختبار أكثر من متغيرين في كل مرحلة فان هذا يسمى بالتحليل متعدد المتغيرات.

* الاحصاء الوصفي والتحليلي.

ان السؤال الثاني الذي يجب ان يُجاب عليه هو معرفة ما اذا كان الشخص الباحث يهتم بوصف العينة أم عمل استنتاجات عن مجتمع الدراسة الذي تم اختيار العينة منه. فالاحصاء الوصفي هو ذلك الجزء من علم الاحصاء الذي يهتم بتزويد الباحث بمقاييس تلخيصيه للبيانات. فهذه المقاييس تزود الباحث باجابات على النوع المقالي من الأسئلة :

(١) ما هو متوسط مثلاً عمر أو دخل عناصر العينة؟

(٢) ما هو التشتت في الأعمار أو الدخل بالعينة؟

(٣) ما هو مستوى العلاقة بين العمر والدخل مثلاً؟

أما بالنسبة للاحصاء الاستقرائي أو التحليلي فهو ذلك النوع من الاحصاء الذي يسمح للشخص الباحث عمل استنتاجات عن مجتمع الدراسة ككل بناءاً على نتائج العينة. ان هذا الجزء من الاحصاء يعتمد على النظرية الاحتمالية. كما يزود هذا الجزء من الاحصاء اجابات على الأسئلة من النوع التالي :

(١) هل متوسط عمر المجتمع هو ٢٥ سنة؟

(٢) هل مستوى العلاقة ما بين العمر والدخل بالنسبة لمجتمع الدراسة أكبر من الصفر؟

(٣) هل متوسط عمر العينة يساوي متوسط عمر المجتمع؟

ان كلا النوعين من الاحصاء سواء كان الوصفي أم الاستقرائي له تطبيقات مهمة في الأبحاث الادارية. لذلك يتوجب على الشخص الاداري الباحث معرفة نوع الاختبار التحليلي ذو الاهتمام.

* مستوى القياس.

ان السؤال الثالث الذي يريد الاجابة هو معرفة نوع المقياس بالنسبة للمتغيرات

المراد تحليلها، فهل تمّ قياسها على مقياس اسمي أم ترتيبي أم مقياس ذو فترة.

فاذا تمكن الشخص الباحث من معرفة عدد المتغيرات المراد تحليلها في كل مرحلة من مراحل التحليل والاهتمام بالاحصاء الوصفي أم التحليلي وميزان القياس للمتغيرات، فهذا يعني انه في وضع يمكنه من اختيار الاختبار الاحصائي المناسب.

ـ الاحصاء الوصفي Descriptive Statistics

يهدف الاحصاء الوصفي الى تزويد مقاييس تلخيصية للبيانات المتوفرة عن عناصر العينة. وغالباً ما يهتم الشخص الباحث بمقاييس النزعة المركزية ومقاييس التشتت من بين مجموعة المقاييس التلخيصية.

ا) مقاييس النزعة المركزية Measures of Central Tendency

يوجد هناك ثلاثة أنواع من مقاييس النزعة المركزية التي غالباً ما تستخدم في الأبحاث الادارية. وهذه المقاييس هي : الوسط والوسيط والمنوال.

١ ـ الوسط الحسابي The Mean

يعتبر الوسط الحسابي مقياساً مناسباً من مقاييس النزعة المركزية بالنسبة للأبحاث الادارية خاصة اذا كان مقياس البيانات هو مقياس فئوي. والوسط الحسابي هو عبارة عن مجموع المشاهدات مقسوماً على حجم العينة. وعليه فان العلاقة الرياضية للوسط هي :

$$\bar{س} = \frac{\sum_{ر=١}^{ن} س_ر}{ن}$$

مثال : أوجد الوسط الحسابي للمشاهدات التالية ٣، ٤، ٥، ٦، ٧ ؟

$\bar{س}$ = [(٣ + ٤ + ٥ + ٦ + ٧) /٥] = ٥

كـمـا يمكن ايجـاد الوسط الحـسـابي من الجـداول الاحـصـائيـة وذلك عن طريق استخدام المعادلة التالية :

$$\bar{س} = \frac{\sum_{ر=1}^{ع} س_ر \times ت_ر}{ن}$$

$س_ر$ = مركز الفئة $_ر$

$ت_ر$ = تكرار الفئة $_ر$

ع = عدد الفئات .

مثال : أوجد الوسط الحسابي للبيانات المجمعة في الجدول التالي

الفئات	التكرار $ت_ر$	مركز الفئة $س_ر$	$س_ر \times ت_ر$
٣٧-٣٩	٧	٣٨	٢٧٦
٤٠-٤٢	٢	٤١	٨٢
٤٣-٤٥	٤	٤٤	١٧٦
٤٦-٤٨	٢	٤٧	٩٤
المجموع	١٥		٦٢٨

$$\bar{س} = (\sum س_ر ت_ر) / \sum ت_ر$$

= ٦٢٨ /١٥

= ٤١٫٨٧

٢ ـ الوسيط The Median

الوسيط هو القيمة التي تتوسط مجموعة من البيانات بعد ترتيبها اما تصاعدي أو تنازلي. فاذا كان عدد البيانات فردياً فان هناك قيمة وحيدة تتوسط البيانات ويطلق عليها اسم الوسيط وتكون رتبة هذه القيمة هي [(ن + ١) /٢]. أما اذا كان عدد البيانات زوجياً فان هناك قيمتان تتوسطان البيانات وتكون رتب هاتين القيمتين هما (ن/٢ و ن/٢ + ١) ويكون الوسيط هو الوسط الحسابي لتلك القيمتين. ويعتبر مقياس الوسيط مقياساً مناسباً اذا كان مقياس البيانات ترتيبي أو مئوي.

ولايجاد الوسيط من الجداول التكرارية، فانه لا بد من ايجاد الفئة الوسيطية. والفئة الوسيطية هي أول فئة يزيد تكرارها التراكمي عن ن/٢ أو يساويها حيث ن هو مجموع التكرارات. وبعد ذلك نجد طول الفئة والحد الأدنى الفعلي للفئة الوسيطية وتكرار الفئة الوسيطية والتكرار المتراكم للفئة قبل الوسيطية. والعلاقة الرياضية التي تستخدم لايجاد الوسيط هي :

$$\text{الوسيط} = أ = \left(\frac{ن/٢ - ن١}{ت} \right) \times ط$$

حيث ان أ = الحد الأدنى الفعلي للفئة الوسيطية.

ن = مجموع التكرارات.

ن١ = التكرار التراكمي للفئة قبل الوسيطية.

ت = التكرار العادي للفئة الوسيطية.

ط = طول الفئة.

مثال : اذا كان الجدول التالي يمثل التوزيع التكراري لنسب الربح لعشرون مؤسسة في سنة ١٩٩٥م. فأوجد الوسيط ؟

الحد الفعلي الأدنى	التكرار التراكمي	عدد المؤسسات	نسبة الربح
٧٫٥	٧	٧	٨-١١
١١٫٥	١٥	٨	١٢-١٥
١٥٫٥	١٨	٣	١٦-١٩
١٩٫٥	١٨	١	٢٠-٢٣
٢٣٫٥	١٩	١	٢٤-٢٧
		٢٠	المجموع

الحل : لايجاد الوسيط لا بد أولاً من ايجاد التكرار التراكمي والحدود الفعلية الدنيا للفئات ومن ثم نجد الفئة الوسيطية والتكرار التراكمي للفئة قبل الوسيطية وطول الفئة والتكرار العادي للفئة الوسيطية ومن ثم نقوم بالتعويض بالعلاقة الرياضية كما يلي :

ن = ٢٠ ، طول الفئة (ط) = ٥

$$\frac{ن}{٢} = \frac{٢٠}{٢} = ١٠$$

∴ الفئة الوسيطية هي الفئة (١٢-١٥)

تكرار الفئة الوسيطية العادي (ت) = ٨

الحد الفعلي الأدنى للفئة الوسيطية (أ) = ١١٫٥

التكرار التراكمي للفئة قبل الوسيطية (ن١) = ٧

$$\therefore \text{الوسيط} = ١١٫٥ + \frac{١٠ - ٧}{٨} \times ٤$$

$$= ١١٫٥ + (٣/٨) \times ٤$$

$$= ١١٫٥ + ١٫٥$$

$$= ١٣$$

ـ استخدامات الوسط والوسيط

يعتبر كل من الوسط والوسيط مقاييس مهمة ومفيدة من مقاييس النزعة المركزية. وفي بعض الحالات يكون مقياس الوسط أفضل من مقياس الوسيط، وفي بعض الحالات الأخرى يكون العكس هو الصحيح. وتعتبر العوامل التالية من أهم المعايير المستخدمة في تحديد أي الاختبارات يجب أن يُستخدم.

أ ـ الحساسية للبيانات المتطرفة :

غالباً ما يعتبر الوسيط أفضل من الوسط اذا كان الوسط يتأثر بشكل كبير بالبيانات المتطرفة. مثال : أوجد الوسط الحسابي للقيم التالية ٢، ٣، ٣، ٤، ٣٨.

الوسط الحسابي = (٢ + ٣ + ٣ + ٤ + ٣٨) /٥

= ٥٠ /٥

= ١٠

لذلك لا يعتبر الوسط الحسابي مقياساً مناسباً لوصف هذه البيانات. ان المقياس الأفضل قد يكون الوسيط والذي في مثالنا الحالي يساوي (٣). ان تأثير الوسيط بالبيانات المتطرفة يكون أقل مقارنة بالوسط والذي بدوره أدى الى زيادة الوسط الحسابي بشكل ملحوظ.

ب ـ الفئات المقفلة والمفتوحة

اذا كان المطلوب وصف المشاهدات الواقعة في فئة مفتوحة، فانه قد يكون من الأفضل استخدام الوسيط لعدم وجود بديل لأن الوسط الحسابي يتطلب معرفة مجموع المشاهدات وعددها. والعكس قد يكون صحيح بالنسبة للفئات المقفلة.

جـ ـ الملائمة الرياضية

يعتبر الوسط الحسابي من أكثر مقاييس النزعة المركزية ملائمة لامتلاكها خصائص رياضية يفتقر لها مقياس الوسيط. فعلى سبيل المثال الوسط الحسابي لمجتمعين مختلفين أو عينتين مختلفتين هو الوسط الحسابي للوسطين. بينما لا يمكن تحديد الوسيط لمجتمعين أو عينتين مختلفتين معاً اذا كان الوسيط معروفاً لكل مجتمع أو عينة.

د ـ مدى الانحراف العيني

غالباً ما تستخدم المقاييس الاحصائية مثل الوسط والوسيط لتقدير الوسط للمجتمع. ان السبب الرئيسي لتفضيل الوسط على الوسيط هو ان الوسط الحسابي غالباً ما يبدو أكثر ثقة في تقدير الوسط الحسابي للمجتمع مقارنة مع الوسيط. بمعنى آخر، ان احتمال ان يكون الوسط الحسابي للعينة اكثر بُعداً عن الوسط الحسابي للمجتمع هو اقل مقارنة مع الوسيط.

٣ ـ المنوال The Mode

ان احد المقاييس الأخرى المستخدمة من مقاييس النزعة المركزية هو المنوال. ويعرّف المنوال للمشاهدات على انه المشاهدة الأكثر تكراراً من بين مجموعة من المشاهدات. أما بالنسبة للمنوال للجداول التكرارية فهو مركز الفئة الأكثر تكراراً.

ـ العلاقة ما بين الوسط والوسيط والمنوال

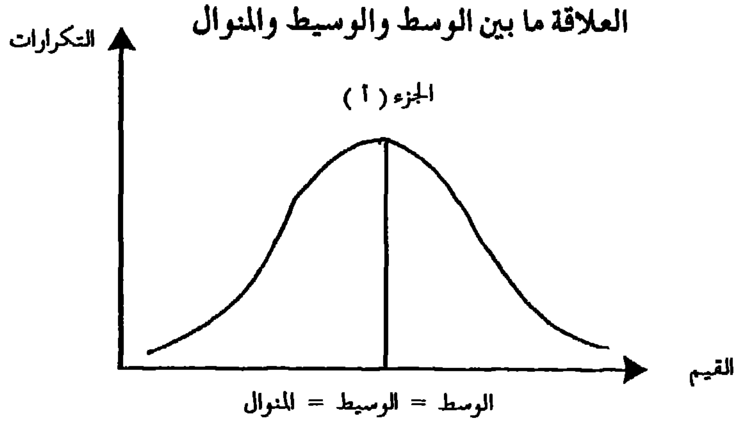
بعد شرح المقاييس الثلاثة الرئيسية (الوسط، الوسيط، المنوال) من مقاييس لنزعة المركزية، فانه لا بد من وصف العلاقة ما بين هذه المقاييس المختلفة. فاذا كان التوزيع التكراري لمجموعة من البيانات له منوال واحد ومتماثل كما في الجزء أ من الشكل (١٠ ـ ١) فان الوسط يساوي الوسيط ويساوي المنوال. والمقصود بالتوزيع المتماثل هو التوزيع الذي يكون فيه الجهة الواقعة على يمين الوسط مطابقة تماماً للجهة الواقعة على يسار الوسط.

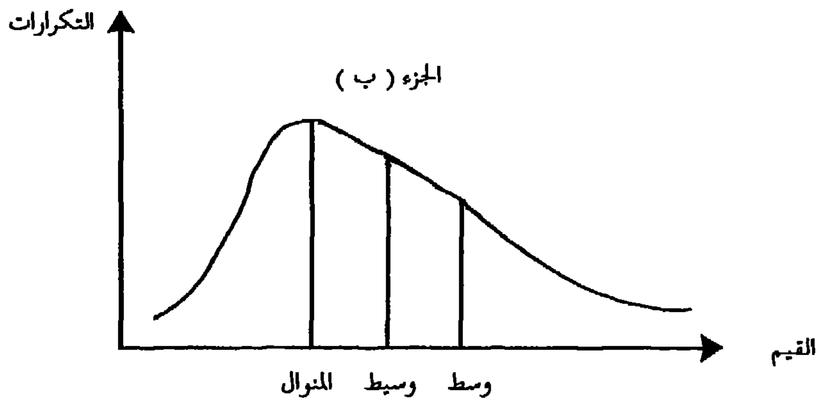
ان معظم التوزيعات التكرارية هي توزيعات تكرارية غير متماثلة، وهذه التوزيعات اما ان تكون مائلة الى جهة اليمين كما في الجزء ب من الشكل (١٠ ـ ١) أو مائلة الى جهة اليسار كما في الجزء جـ من الشكل (١٠ ـ ١). فالتوزيع التكراري المائل الى جهة اليمين يكون له ذيل طويل في جهة اليمين، أما التوزيع التكراري المائل الى جهة اليسار يكون له ذيل طويل في جهة اليسار.

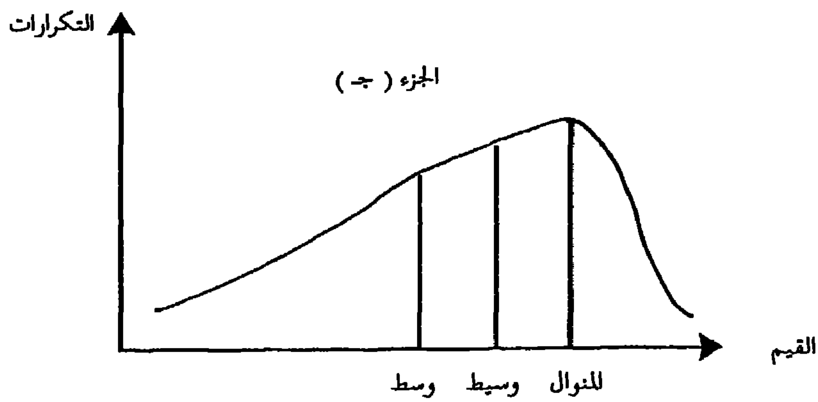
وكما يظهر في الشكل (١٠ ـ ١)، اذا كان التوزيع التكراري مائل الى جهة اليمين فانه غالباً ما يكون الوسط أكبر من الوسيط وأكبر من المنوال. أما اذا كان التوزيع التكراري مائل الى جهة اليسار فانه غالباً ما يكون المنوال أكبر من الوسيط وأكبر من الوسط.

شكل (١٠ ـ ١)

العلاقة ما بين الوسط والوسيط والمنوال







ب) مقاييس التشتت Measures of Dispersion

ان مقاييس النزعة المركزية لا تستطيع تزويد الباحث بالمعلومات الكافية لفهم التوزيع المراد اختياره بشكل كامل. فعلى سبيل المثال الأرقام ١٠، ١٥، ٢٠ والأرقام ٥، ١٠، ٣٠ لها نفس الوسط الحسابي (١٥). ان هذا يعني وبوضوح ان الوسط الحسابي وحده لا يكون كافياً للحصول على وصف جيد للبيانات. لذلك فانه لا بد من الاهتمام بمدى انتشار التوزيع. بمعنى آخر فانه لا بد من الحاجة الى مقياس يستطيع قياس انتشار التوزيع بالنسبة للمتغير. ويطلق على هذا النوع من المقاييس بمقاييس التشتت. ومن هذه المقاييس :

١ ـ المدى Range

يعرّف المدى للبيانات على انه الفرق ما بين أعلى قيمة واصغر قيمة. فاذا كان المدى صغيراً فاننا نستنتج بان البيانات محصورة في مسافة قصيرة، واذا كان المدى كبيراً فان هذا يعني ان البيانات تقع ضمن مسافة كبيرة.

ويعرّف المدى للبيانات المجمعة أو التوزيعات المركزية على انه الفرق ما بين الحد الأعلى للفئة العليا والحد الأدنى للفئة الدنيا. ومن تعريف المدى، يتبين لنا انه لا يعتمد على جميع البيانات ولكن يعتمد على أكبر قيمة وأدنى قيمة فقط. وهذا ما يقلل من اهمية المدى خاصة اذا كانت القيمتين المتطرفتين (أكبر قيمة وأدنى قيمة) قيمتان شاذتان، ففي هذه الحالة يكون المدى كبيراً بينما مفردات البيانات ليست متباعدة عن بعضهما البعض.

مثال : اذا كانت علامات طلاب مادة أساليب البحث العلمي كما يلي :

١٠٠، ٧١، ٧٢، ٦٥، ٦٠، ٧٠، ٦٨، ٦٥، ٦٦، ٢٠

فان المدى = ١٠٠ ـــ ٢٠ = ٨٠

بينما معظم العلاقات واقعة ما بين ٧٢ و ٦٠، اي ان العلامات متقاربة جداً بينما المدى يظهر عكس ذلك.

٢ ـ الانحراف المتوسط The Mean Deviation

ان أحد مقاييس التشتت التي تخطر على البال هو مجموع انحرافات البيانات عن وسطها الحسابي $\sum$ (س ر − سَ). ولكن هذا المجموع يساوي دائماً صفراً لأن مجموع الانحرافات الموجبة عن الوسط الحسابي يساوي مجموع الانحرافات السالبة. لذلك لا بد من التخلص من الاشارة السالبة حتى نحصل على مقياس ذي معنى. ان إحدى الطرق المستخدمة في التخلص من الاشارة السالبة هي عن طريق اخذ القيمة المطلقة (Absolute Value) والتي تعرّف بالمعادلة :

$$|س| = \begin{cases} س & س \geq ٠ \\ -س & س < ٠ \end{cases}$$

وباستخدام القيمة المطلقة نحصل على تعريف الانحراف المتوسط وكما يلي :

$$\text{الانحراف المتوسط} = \frac{\sum_{ر=١}^{ن} |س_ر - \bar{س}|}{ن}$$

حيث ان سَ هو الوسط الحسابي للبيانات.

مثال : أوجد الانحراف المتوسط للبيانات ٧، ٩، ٤، ٥، ٨، ١٠، ٧، ٦.

الحل : اولاً نجد الوسط الحسابي سَ

$$\bar{س} = \sum س_ر / ن$$

$$\bar{س} = (٧ + ٩ + ٤ + ٥ + ٨ + ١٠ + ٧ + ٦) / ٨$$

$$\bar{س} = ٥٦ / ٨$$

$$\bar{س} = ٧$$

$$\text{اذن الانحراف المتوسط} = \frac{\sum |س_ر - \bar{س}|}{ن}$$

$$= \frac{|٧-٧| + |٩-٧| + |٤-٧| + |٥-٧| + |٨-٧| + |١٠-٧| + |٧-٧| + |٦-٧|}{٨}$$

$$= \frac{٠ + ٢ + ٣ + ٢ + ١ + ٣ + ٠ + ١}{٨}$$

$$= \frac{١٢}{٨}$$

$$= ١ر٥$$

أما في حالة التوزيعات التكرارية فاننا نستخدم التعريف التالي للانحراف المتوسط

$$\text{الانحراف المتوسط} = \frac{\sum |س_ر - \bar{س}| \times ت_ر}{ن}$$

حيث ان : $\bar{س}$ الوسط الحسابي

$ت_ر$ التكرار

ن مجموع التكرارات

مثال : احسب الانحراف المتوسط للتوزيع التكراري في الجدول التالي :

$س_ر$	$ت_ر$
٣	٢
٧	٥
١١	٩
١٥	٧
١٩	٢

الحل : نرتب الحسابات في الجدول التالي

$س_ر$	$ت_ر$	$س_ر$ X $ت_ر$	$س_ر$ – سَ	|$س_ر$ – سَ|	|$س_ر$ – سَ| X $ت_ر$
٣	٢	٦	–٨ر٣٢	٨ر٣٢	١٦ر٦٤
٧	٥	٣٥	– ٤ر٣٢	٤ر٣٢	٢١ر٦٠
١١	٩	٩٩	– ر٣٢	ر٣٢	٢ر٨٨
١٥	٧	١٠٥	٣ر٦٨	٣ر٦٨	٢٥ر٧٦
١٩	٢	٣٨	٧ر٦٨	٧ر٦٨	١٥ر٣٦
المجموع	**٢٥**	**٢٨٣**			**٨٢ر٢٤**

الوسط الحسابي = $\frac{٢٨٣}{٢٥}$

= ١١ر٣٢

$$\text{اذن الانحراف المتوسط} = \frac{٨٢٫٢٤}{٢٥}$$

$$= ٣٫٢٩$$

يعتمد الانحراف المتوسط على جميع مفردات البيانات وهو سهل التعريف وسهل الحساب الا انه لا يخضع للعمليات الجبرية بسهولة حيث يجب تعديل الاشارة ويجب معرفة المفردات بعينها اذا ما أردنا حساب قيمته. ويتضح هذا من عدم وجود طريقة جبرية لحساب الانحراف المتوسط للمجموعة الناتجة عن دمج مجموعتين من البيانات اذا علم عدد مفردات كل منها ووسطها الحسابي وانحرافها المتوسط. ففهي هذه الحالة يجب معرفة جميع المفردات لنتمكن من حساب انحرافها المتوسط.

٣ ـ الانحراف المعياري والتباين The Standard Deviation and Varience

ان الطريقة الثانية للتخلص من الاشارة السالبة للانحرافات عن الوسط الحسابي هي بتربيع تلك الانحرافات واستعمالها في حساب التباين والذي جذره التربيعي يساوي الانحراف المعياري. وعلى الرغم من استخدام المدى والانحراف المتوسط لقياس التشتت في بعض الأحيان، الا ان التباين والانحراف المعياري من أكثر المقاييس أهمية في قياس متوسط التشتت. ويرمز لمقياس التباين بالنسبة لمجتمع الدراسة بالرمز (σ^2) والذي يعرّف بمجموع مربع الانحرافات بالنسبة للمشاهدات عن وسطها. ويعرّف رياضياً كما يلي :

$$\sigma^{٢} = \frac{\sum_{ر=١}^{ن} (س_ر - \mu)^{٢}}{ن}$$

أما بالنسبة للتباين بالنسبة للعينة فيرمز له بالرمز (S^2) ويعرف رياضياً كما يلي :

$$S^2 = \frac{\sum (\text{س}_{\text{ر}} - \bar{\text{س}})^2}{\text{ن} - ١}$$

أما بالنسبة للانحراف المعياري بالنسبة للمجتمع الدراسي فهو عبارة عن الجذر التربيعي للتباين وكذلك نفس الشيء بالنسبة للانحراف المعياري بالنسبة للعينة. وعليه يكون الانحراف المعياري لمجتمع الدراسة كما يلي :

$$\sigma = \sqrt{\frac{\sum (\text{س}_{\text{ر}} - \mu)^2}{\text{ن}}}$$

بينما الانحراف المعياري بالنسبة للعينة فيعرّف كما يلي :

$$S = \sqrt{\frac{\sum (\text{س}_{\text{ر}} - \bar{\text{س}})^2}{\text{ن} - ١}}$$

مثال : أوجد الانحراف المعياري للمشاهدات التالية :

١٩ر٥، ١٦ر٨، ١٦ر٦، ١٢ر٣، ١٨ر٧، ١٨، ١٨ر٤، ١٩ر٣، ٢٥ر٢، ١٦، ١٩ر٢، ٢١ر٥، ٩ر٣، ١٨، ١٦ر٦، ٣٢ر٦، ٢٠ر١، ١٤ر١، ٢٠ر٥، ١٦ر٨، ١٣ر٧.

الحل : نرتب الحسابات في الجدول التالي

س ر	س ر − سَ	(س ر − سَ)²
١٩ر٥	١ر٢٥	١ر٥٦٢٥
١٦ر٨	− ١ر٤٥	٢ر١٠٢٥
١٦ر٦	− ١ر٦٥	٢ر٧٢٢٥
١٢ر٣	− ٥ر٩٥	٣٥ر٤٠٢٥
١٨ر٧	ر٤٥	٠ر٢٠٢٥
١٨	− ر٢٥	٠ر٠٦٢٥
١٨ر٤	ر١٥	٠ر٠٢٢٥
١٩ر٣	١ر٠٥	١ر١٠٢٥
٢٥ر٢	٦ر٩٥	٤٨ر٣٠٢٥
١٦	− ٢ر٢٥	٥ر٠٦٢٥
١٩ر٢	ر٩٥	٠ر٩٠٢٥
٢١ر٥	٣ر٢٥	١٠ر٥٦٢٥
٩ر٣	− ٨ر٩٥	٨٠ر١٠٢٥
١٨	− ر٢٥	٠ر٠٦٢٥
١٦ر٦	− ١ر٦٥	٢ر٧٢٢٥
٣٢ر٦	١٤ر٣٥	٢٠٥ر٩٢٢٥
٢٠ر١	١ر٨٥	٣ر٤٢٢٥
١٤ر١	− ٤ر١٥	١٧ر٢٢٢٥
٢٠ر٥	٢ر٢٥	٥ر٠٦٢٥
١٦ر٨	− ١ر٤٥	٢ر١٠٢٥
١٣ر٧	− ٤ر٥٥	٢٠ر٧٠٢٥
٢٨٣ر٢		٤٤٥ر٣٣٢٥

$$\text{الوسط الحسابي} = \frac{٢٨٣٫٢}{٢١}$$

$$= ١٨٫٢٥$$

$$\text{اذن الانحراف المعياري} = \sqrt{\frac{\sum (س_ر - \bar{س})^٢}{ن - ١}}$$

$$S = \sqrt{\frac{٤٤٥٫٣٣٢٥}{٢١ - ١}}$$

$$S = ٤٫٧٢$$

اما بالنسبة للتوزيعات التكرارية فاننا نستخدم التعاريف الرياضية التالية لحساب الانحراف المعياري بالنسبة لمجتمع الدراسة والعينة.

$$\sigma = \sqrt{\frac{\sum (س_ر - \mu)^٢ \times ت_ر}{ن}}$$

اما $$S = \sqrt{\frac{\sum (س_ر - \bar{س})^٢ \times ت_ر}{ن - ١}}$$

مثال : أوجد الانحراف المعياري لنسب الربح للمؤسسات التالية

نسبة الربح	عدد المؤسسات
٨-١١	٧
١٢-١٥	٨
١٦-١٩	٣
٢٠-٢٣	١
٢٤-٢٧	١

الحل : نرتب الحسابات في الجدول التالي

نسبة الربح	عدد المؤسسات ت ر	مركز الفئة س ر	س ر X سَ	(س ر – سَ)٢	(س ر – سَ)٢	(س ر – سَ)٢ X ت ر
٨-١١	٧	٩ر٥	٦٦ر٥	-٢ر٤	١٧ر٦٤	١٢٣ر٤٨
١٢-١٥	٨	١٣ر٥	١٠٨	-ر٢	ر٠٤	ر٣٢
١٦-١٩	٣	١٧ر٥	٥٢ر٥	٣ر٨	١٤ر٤٤	٤٣ر٣٢
٢٠-٢٣	١	٢١ر٥	٢١ر٥	٧ر٨	٦٠ر٨٤	٦٠ر٨٤
٢٤-٢٧	١	٢٥ر٥	٢٥ر٥	١١ر٨	١٣٩ر٢٤	١٣٩ر٢٤
المجموع	٢٠		٢٧٤			٣٦٧ر٢

الوسط الحسابي (سَ) = $\frac{٢٧٤}{٢٠}$ = ١٣ر٧

∴ الانحراف المعياري = $\sqrt{\frac{٣٦٧ر٢}{١٩}}$

$$S = \sqrt{١٩ر٣٣}$$

$$S = ٤ر٤٠$$

ويمكن استخدام التعريف التالي للنحراف المعياري للتسهيل

$$S = \sqrt{\frac{\sum \text{س}_{\text{ر}}^{٢} \times \text{ت}_{\text{ر}} - \frac{١}{\text{ن}} \left(\sum \text{س}_{\text{ر}} \text{ت}_{\text{ر}} \right)^{٢}}{\text{ن} - ١}}$$

وحل المثال اعلاه باستخدام هذه العلاقة نجد ما يلي :

(١) $\sum \text{س}_{\text{ر}}^{٢} \times \text{ت}_{\text{ر}}$ = (٩ر٥)^{٢} × ٧ + (١٣ر٥)^{٢} × ٨ + (١٧ر٥)^{٢} × ٣ + (٢١ر٥)^{٢} × ١ + (٢٥ر٥)^{٢} × ١

= ٦٣١ر٧٥ + ١٤٥٨ + ٩١٨ر٧٥ + ٤٦٢ر٢٥ + ٦٥٠ر٢٥

= ٤١٢١

(٢) $\sum \text{س}_{\text{ر}} \times \text{ت}_{\text{ر}}$ = (٩ر٥) × ٧ + (١٣ر٥) × ٨ + (١٧ر٥) × ٣ + (٢١ر٥) × ١ + (٢٥ر٥) × ١

= ٦٦ر٥ + ١٠٨ + ٥٢ر٥ + ٢١ر٥ + ٢٥ر٥

= ٢٧٤

اذن الانحراف المعياري هو

$$S = \sqrt{\frac{٤١٢١ - \frac{١}{٢٠}(٢٧٤)^{٢}}{٢٠ - ١}}$$

$$S = \sqrt{\frac{٤١٢١ - ٣٧٥٣{,}٨}{١٩}}$$

$$S = \sqrt{\frac{٣٦٧{,}٢}{١٩}}$$

$$S = \sqrt{١٩{,}٣٣}$$

$$S = ٤{,}٤$$

وعليه نلاحظ ان الانحراف المعياري باستخدام كلا المعادلتين كان نفس الشيء. ويجب ان تكون جميع الاجابات بالنسبة للمعادلتين متطابقة في جميع المجالات.

- تفسير الانحراف المعياري Interpretation of the Standerd Deviation

يعتبر مقياس الانحراف المعياري من أكثر مقاييس التشتت أهمية. فاذا كان التوزيع التكراري للمجتمع مطابقاً لما يدعى بالتوزيع الطبيعي، فاننا نستطيع معرفة الحالات من المجتمع والتي تقع ضمن انحراف معياري واحد أو اثنين أو ثلاثة انحرافات معيارية عن وسط المجتمع. وبالتحديد فانه يقع ٦٨٫٣٪ من حالات مجتمع الدراسة ما بين ± ١ انحراف معياري عن الوسط، و ٩٥٫٤٪ من الحالات تقع ما بين ± ٢ انحرافين معيارين عن الوسط و ٩٩٫٧٪ من الحالات تقع ما بين ± ٣ انحرافات معيارية عن الوسط.

لذلك، اذا كنا نعلم ان قطر الأنابيب المنتجة في شركة ما مطابقاً لما يدعى بالتوزيع الطبيعي أو المتماثل وكان وسطه الحسابي ١٠ سم والانحراف المعياري هو ١ سم، فان هذا يعني ان ٣ر٦٨ من الأنابيب المنتجة في هذه الشركة سوف يكون اقطارها واقعة ما بين ٩ سم و ١١ سم، و ٤ر٩٥٪ من الأنابيب سوف تكون اقطارها واقعة ما بين ٨ سم و ١٢ سم و ٧ر٩٩٪ من الأنابيب سوف تكون اقطارها واقعة ما بين ٧ سم و ١٣ سم.

لذلك تعتبر هذه المعلومات مهمة ومفيدة. فعلى سبيل المثال : اذا كان يجب ان تقع الأقطار للأنابيب المنتجة ما بين ٩ سم و ١١ سم لتحقق الشروط، فان هذا يعني ان نسبة الوحدات غير المقبولة هي ٦ر٤٪ .

انه من الضروري ادراك هنا ان هذه النتائج تنطبق على المجتمعات التي يكون توزيعها التكراري توزيعاً طبيعياً. أما اذا كان التوزيع التكراري للمجتمع هو توزيع غير طبيعي، فانه أيضاً يمكن عمل استنتاجات عن نسبة الحالات التي تقع ما بين انحرافات معيارية محددة بناءاً على عدم مساواة كبيكيف (Chebyshev's Inequality) . وللحصول على معلومات اضافية عن هذا الموضوع فانه يمكن الرجوع الى كتب الاحصاء.

أسئلة

س١ : ما هو المقصود بتحليل البيانات الأحادي، الثنائي، المتعدّد ؟

س٢ : قارن ما بين الاحصاء الوصفي والاحصاء التحليلي ؟

س٣ : لماذا يكون من المهم وصف مجموعة من البيانات بناءاً على مقاييس النزعة المركزية ومقاييس التشتت ؟

س٤ : قارن ما بين الطرق المختلفة التي تستخدم في عرض البيانات ؟

س٥ : اذا كانت البيانات التالية تمثل أعداد الطلبة في جامعة مؤتة الجناح المدني خلال السنوات ١٩٩٠ و ١٩٩٤.

السنة	الذكور	الاناث
١٩٩٠	٦٠٠	٤٠٠
١٩٩١	٩٠٠	٦٠٠
١٩٩٢	١٣٠٠	٧٠٠
١٩٩٣	٢٠٠٠	٩٠٠
١٩٩٤	٢٢٠٠	١١٠٠

المطلوب : عرض البيانات بيانياً باستخدام الطرق المختلفة.

س٦ : اذا كان الجدول التالي يمثل التوزيع التكراري للأجور الأسبوعية لـ ٦٠٠ عامل في شركة ما

الأجرة الأسبوعية	عدد العمال
١٦-٢٠	٧
٢١-٢٥	١٧
٢٦-٣٠	١١٦
٣١-٣٥	٢٠٥
٣٦-٤٠	١٣٤
٤١-٤٥	٦٠
٤٦-٥٠	٣٣
٥١-٥٥	١٧
٥٦-٦٠	١١

أوجد ما يلي :

أ - الوسط الحسابي للأجور.

ب - الوسيط.

جـ - المنوال.

د - الانحراف المتوسط.

هـ - الانحراف المعياري.

و - المدى.

س٧ : اذا كان لديك البيانات التالية

٩، ١٠، ٧، ١٢، ١٥، ٤، ٨

فاوجد ما يلي : أ- الوسيط.

ب - الانحراف المتوسط.

جـ- الانحراف المعياري.

س٨ : أذكر المعايير المستخدمة في تحديد اي المقاييس التي يجب استخدامها من مقاييس النزعة المركزية ؟

الفصل الحادي عشر

اختبار الفرضيات

Hypothesis Testing

اختبار الفرضيات
Hypothesis Testing

ـ مقدمة Introduction

يلجأ الباحثون في كثير من الحالات الى استخدام العينة لمعرفة خصائص مجتمع الدراسة للمساعدة في اتخاذ القرارات بناءاً على بيانات العينة. وعلى الرغم من عدم تطابق نتائج العينة الى حد كبير مع نتائج مجتمع الدراسة، الا ان هذا الاجراء يعتبر مقبولاً لمعظم الاحصائيون. لذلك يهتم اختبار الفرضيات بمعرفة ما اذا كان هناك اختلاف هام ما بين نتائج دراسة العينة ومعالم مجتمع الدراسة عن طريق استخدام الاختبارات الاحصائية. وبما ان الاحصائيون غالباً ما يستخدموا اختبار الفرضيات، وان الطرق التي يستخدموها والنتائج التي يحصلوا عليها تلعب دوراً مهماً في الادارة والاقتصاد، فانه من المهم جداً فهم المبادىء والمفاهيم التي يتضمنها اختبار الفرضيات وكيفية تطبيق هذه المفاهيم والمبادىء في حل المشاكل العملية.

والفرضية الاحصائية هي عبارة عن جملة حول مجتمع احصائي أو أكثر من مجتمع بحيث تدور هذه الجملة بالغالب حول معالم المجتمع الاحصائي. فمثلاً يمكن لمعلّم ان يضع فرضية حول نسبة الطلبة المتميزين في إحدى الصفوف بقوله ان نسبتهم ١٠٪.

ـ مفهوم الفرضية الأساسية The Concept of a Null Hypothesis

يبدأ اختبار الفرضيات بجملة تدعى الفرضية الأساسية. والفرضية الأساسية غالباً ما تفترض ان معالم مجتمع الدراسة لها قيمة معينة أو مجموعة من القيم. فعلى سبيل

المثال، قد يرغب شخص ما في اختبار ما اذا كان متوسط عمر طلاب صف ما هو ٢٠ سنة. ان شكل الفرضية الأساسية (ف٠) هو :

$ف_٠ : \mu = ٢٠$

ان الفرضية الأساسية هي الفرضية التي نقوم بتقيمها في اختبار الفرضيات. فالفرضية الأساسية في بعض الحالات يمكن أن ترفض ويتم قبول الفرضية البديلة ($ف_١$)، أو نقوم

بقبول الفرضية الأساسية بعد اجراء الاختبار. وفي هذه الحالة، فاننا لا نستطيع استنتاج ان الفرضية الأساسية هي صالحة انما كل ما يمكن قوله هو عدم وجود دليل لرفض الفرضية الأساسية. إنه من الممكن جداً اثبات عدم صحة الفرضية التي تمّ قبولها في وقت لاحق بعد تجميع بيانات جديدة عن العينة. انه لمن المهم جداً أن نلاحظ اننا نقوم باختبار قيمة معلمة مجتمع الدراسة بناءاً على البيانات المجمعّة عن العينة، لذلك فانه من الممكن جداً لقيمة أو قيم العينة أن تختلف عن قيم معالم مجتمع الدراسة بسبب أخطاء العينة. وباختبار التوزيع العيني نستطيع أن نحدد ما اذا كانت قيمة العينة تختلف اختلاف كافٍ عن قيمة الفرضية والذي يحصل بناءاً على أخطاء ناتجة عن العينة. فاذا كان الاختلاف أكبر من الخطأ الناتج عن الخطأ العيني، فاننا سوف نقوم برفض الفرضية الأساسية.

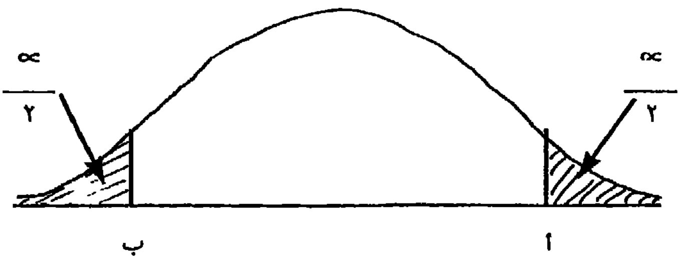
ان وضع الفرضية بالشكل الذي وضعناه أعلاه يتضمن عدد من الفرضيات البديلة. ومن الجدير بالذكر ان الفرضية البديلة (ف١) لها ثلاثة أشكال هي :

$ف_٠ : \mu = ٢٠$

$ف_١ : \mu \neq ٢٠$

ففي مثل هذه الحالة نقوم برفض الفرضية الأساسية اذا كان متوسط أعمار الطلبة اما أكبر أو أقل من (٢٠) سنة. لهذا فان اهتمام الشخص الباحث يكون حول جهتي

التوزيع بالنسبة للتوزيع العيني لمتوسط الأعمار، وعليه فاننا نقوم بوضع نصف قيمة (∝) في كل طرف من طرفي توزيع دالة الاختبار. فاذا مثل الشكل (١١ - ١) منحنى توزيع دالة الاختبار



الشكل (١١ - ١)

فاننا نقوم بايجاد قيمة كل من أ و ب بحيث يكون

احتمال دالة الاختبار فوق أ = دالة الاختبار تحت ب = $\frac{\propto}{٢}$

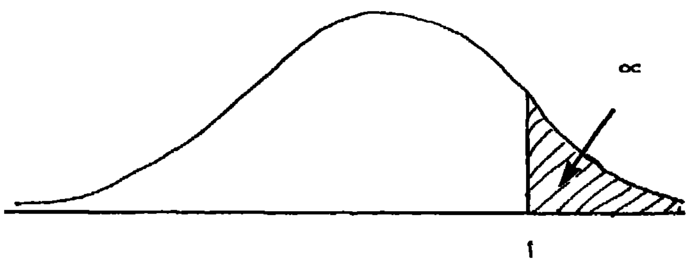
وعليه، فاننا نقوم بقبول الفرضية الأساسية اذا وقعت قيمة الاختبار المحسوبة من العينة ما بين العددين أ و ب ونرفض الفرضية الأساسية اذا وقعت القيمة المحسوبة للاختبار في المنطقة المضللة والتي تمثل منطقة الرفض. ويطلق على هذا النوع من الاختبارات بالاختبارات ذو الطرفين «Two-Tailed» لاننا نقوم باختبار نهايتي التوزيع العيني.

٢ ـ الفرضية البديلة ذات الذيل الأعلى مثل

ف. : μ = ٢٠

ف١ : μ > ٢٠

ففي مثل هذه الحالة، فان اهتمام الباحث يتركز حول الطرف الأعلى من التوزيع العيني لمتوسط الأعمار. وعليه فان الشخص الباحث يقوم برفض الفرضية الأساسية اذا كان متوسط اعمار الطلبة أكبر من (٢٠) سنة. لذلك، فان الشخص الباحث في مثل هذه الحالة يقوم بوضع كل قيمة ($\propto$) في الطرف الأعلى من التوزيع العيني للاختبار. فاذا مثل الشكل (١١ ـ ٢) منحنى توزيع دالة الاختبار.



الشكل (١١ ـ ٢)

فاننا نقوم بايجاد قيمة أ بحيث يكون

احتمال دالة الاختبار فوق أ = $\propto$

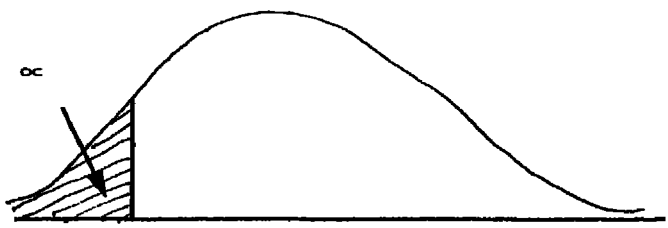
وعليه، فاننا نقوم بقبول الفرضية الأساسية اذا كانت القيمة المحسوبة للاختبار اقل من أ، ونرفض الفرضية الأساسية اذا كانت القيمة المحسوبة للاختبار أكبر من أ. ويطلق على هذا النوع من الاختبارات بالاختبارات ذو الطرف الواحد One-Tailed Test، لأننا نقوم باختبار التوزيع العيني من جهة واحدة.

٣ـ الفرضية البديلة ذات البديل الأدنى مثل

ف٠ : μ = ٢٠

ف١ : μ > ٢٠

ففي مثل هذه الحالة، فان اهتمام الباحث يتركز حول الطرف الأدنى من التوزيع لمتوسط الأعمار. وعليه فان الشخص الباحث يقوم برفض الفرضية الأساسية اذا كان متوسط اعمار الطلبة أقل من (٢٠) سنة. لذلك فان الشخص الباحث في مثل هذه الحالة يقوم بوضع كل قيمة ($\propto$) في الطرف الأدنى من التوزيع العيني للاختبار. فاذا مثل الشكل (١١ ـ ٣) منحنى توزيع دالة الاختبار.



الشكل (١١ ـ ٣)

فاننا نقوم بايجاد قيمة ب بحيث يكون

احتمال دالة الاختبار تحت ب = $\propto$

وعليه، فاننا نقوم بقبول الفرضية الأسياسية اذا كانت القيمة المحسوبة للاختبار أكبر من (ب)، ونرفض الفرضية الأساسية اذا كانت القيمة المحسوبة للاختبار أقل من

(ب) . ويطلق على هذا النوع من الاختبارات بالاختبارات ذو الطرف الواحد One-، Tailde Test، لاننا نقوم باختبار التوزيع العيني من جهة واحدة.

ـ الأخطاء الممكنة Possible Errors

في اختبار الفرضيات، اما أن يقوم صاحب القرار (Decision Maker) بقبول (Accpet) الفرضية الأساسية (ف٠) أو رفضها (Reject). إن القرار اما أن يكون صائب أو خاطيء. فاذا كانت الفرضية الأساسية صحيحة وتمّ قبولها، فان القرار صحيح. أما اذا كانت الفرضية الأساسية خاطئة وتمّ قبولها، فان القرار خاطيء. وعليه فان رفض الفرضية الأساسية الصحيحة وعدم رفض الفرضية الأساسية الخاطئة يشير الى وقوع خطأ في القرار. ويطلق على هذه الاخطاء بالنوع الأول (Type I) والنوع الثاني (Type II) من الأخطاء على التوالي. كما تسمى هذه الأخطاء احياناً بخطأ (∝) وخطأ (ß) على التوالي. والجدول (١١ ـ ١) يمثل ملخص لنتائج العينة واحتمالاتها المتوقعة.

جدول (١١ ـ ١)

ملخص لأخطاء اختبار الفرضيات

نتائج العينة	ظروف صحيحة	
	ف صحيحة	ف٠ خاطئة
لا ترفض ف٠	(١) القرار صحيح (٢) الاحتمال = ١ ـ ∝	(١) خطأ من النوع الثاني (٢) الاحتمال = ß
ارفض ف٠	(١) خطأ من النوع الأول (٢) الاحتمال = ∝	(١) القرار صحيح (٢) الاحتمال = ١ ـ ß

ـ خطوات اختبار الفرضيات Steps in Hypothesis Testing

ان الخطوات التي سوف نستخدمها في اختبار الفرضيات هي :

١ ـ تشكيل الفرضية الأساسية والبديلة.

٢ ـ اختيار الاختبار الاحصائي المناسب بناءاً على نوع البيانات المتوفرة.

٣ ـ تحديد مستوى الثقة (Significance Level) ، ∝

٤ ـ ايجاد القيمة المعيارية للاختبار الاحصائي من الجداول المخصصة بناءاً على قيم معطاة لـ (∝) . Critical Value

٥ ـ إيجاد القيمة الاحصائية للاختبار الذي تمّ اختباره في الخطوة الثانية Statistical Value

٦ ـ مقارنة القيمة الاحصائية للخطوة الخامسة مع القيمة المعيارية للخطوة الرابعة. اذا كانت القيمة الاحصائية اكبر من القيمة المعيارية، فاننا سوف نقوم برفض الفرضية الأسياسية لأن القيمة الاحصائية تكون قد وقعت بعيداً عن التوزيع الذي نعتبره جزءاً من التوزيع العيني للفرضية الأساسية. أما اذا كانت القيمة الاحصائية أقل من القيمة المعيارية، فاقبل الفرضية الأساسية.

ـ اختبار الوسط لعينة واحدة One- Sanple Test of a Mean

فبعد أن تم وصف المبادىء الأساسية للنظرية الاحصائية في اختبار الفرضيات، فانه بالامكان الآن تفصيل اهم الاختبارات الاحصائية. وفي هذا الجزء، فاننا سوف نهتم بالحالات التي تكون فيها البيانات ذات علاقة بعينة واحدة. بمعنى آخر، سوف يهتم هذا الجزء بشرح الاختبارات.الاحصائية لوسط المجتمع الحسابي بناءاً على العينات العشوائية البسيطة. ان هذا النوع من الاختبارات معروف ويشكل الأساس للاختبارات الاحصائية الأخرى.

يوجد هناك اختباران معروفان ومناسبان لهذا الغرض هما اختبار Z (Z Test) واختبار t (t Test). ان استخدام أحد هذين الاختبارين يعتمد على معرفة الشخص بالانحراف المعياري لمجتمع الدراسة (σ) وحجم العينة المستخدم.

ـ اختبار Z

يستخدم اختبار Z لمقارنة الوسط الحسابي الذي يتم استخراجه من عينة مع وسط حسابي مقترح لمجتمع دراسي، ومن ثم اتخاذ القرار فيما اذا كان الوسط الحسابي للعينة يسمح باستنتاج ان الوسط الحسابي المقترح لمجتمع الدراسة هو صحيح. ويعتبر هذا الاختبار مناسباً في الحالات التالية.

١ ـ اذا كان حجم العينة أي حجم والانحراف المعياري لمجتمع معروف (σ).

٢ ـ اذا كان حجم العينة أكبر من (٣٠) والانحراف المعياري للمجتمع (σ) مجهول.

وفي الحالات التي يكون فيها حجم العينة أقل من (٣٠) والانحراف المعياري لمجتمع الدراسة (σ) مجهول، فانه يجب استخدام اختبار (t) والذي سوف يتم شرحه فيما بعد.

دعنا الآن نقوم بتوضيح اختبار (Z) من خلال المثال التالي :

يريد طالب عمل دراسة لمعرفة متوسط عمر طلبة جامعة مؤتة. فاختار عينة عشوائية مكونة من (١٠٠) شخص، فوجد ان متوسط عمر العينة (٢٤) سنة والانحراف المعياري (٥). فهل يستطيع الطالب اعتبار متوسط عمر طلبة جامعة مؤتة (٢٣) سنة بالاعتماد على متوسط عمر العينة، عند مستوى الدلالة (∝) ٥٪.

الخطوة الأولى : تشكل الفرضية الأساسية والبديلة. ان شكل الفرضية الأساسية والبديلة في مثالنا هذا هما :

الفرضية الأساسية (ف$_0$) : μ = ٢٣

الفرضية البديلة (ف$_1$) : $\mu \neq$ ٢٣

ان الطالب يريد معرفة ما اذا كان متوسط عمر العينة (٢٤ سنة) يسمح بالنتيجة التي تعتبر ان متوسط عمر مجتمع الدراسة هو (٢٣ سنة). ان الفرضية البديلة شكلت بطريقة بحيث اذا كانت قيمة العينة بعيدة عن العمر (٢٣) سنة من كلا الطرفين (أعلى أو أسفل)، فاننا سوف نقوم برفض الفرضية الأساسية. لذلك، فاننا سوف نستخدم اختبار ذو الطرفين.

اما اذا كانت الفرضية البديلة ف$_1$: $\mu <$ ٢٣، فاننا سوف نرفض الفرضية الأساسية في الحالة التي يكون فيها متوسط العمر أقل من الفرضية الأساسية (ف$_0$). ويطلق على هذا النوع من الاختبارات بالاختبارات ذو الطرف الواحد.

انه لمن الضروري جداً لأي شخص معرفة أي الاختبارات التي يتعامل معها، أهو اختبار ذو طرفين أم اختبار ذو طرف واحد، لأن عدد الأطراف أو الذيول يؤثر على القيمة المعيارية التي يجب اخراجها من الجداول الاحصائية المخصصة لكل اختبار.

الخطوة الثانية : اختيار الاختبار الاحصائي الأمثل. ان الاختبار الاحصائي الأمثل هو اختبار Z لأن حجم العينة أكبر من (٣٠). كما ان هذا الاختبار يعتمد على طبيعة التوزيع العيني للوسط. اننا نعلم من خلال نظرية الحدود المركزية (Central Limit Theorem) ان الوسط الحسابي الذي يتم حسابه عن طريق العينة يأتي من التوزيع العيني للأوساط والتي تشكل منحنى طبيعي. كما ان المساحة تحت المنحنى الطبيعي وعدد الأوساط الحسابية الموزعة ما بين عدد محدود من الأخطاء المعيارية للوسط لهذا التوزيع معروفة. لهذا، فاننا نستطيع تحديد احتمال أي وسط عيني يأتي من التوزيع العيني للأوساط والذي يقع حول وسط مقترح للمجتمع.

الخطوة الثالثة : تحديد مستوى الثقة. يعتبر مستوى الثقة أو الدلالة المقياس الاحتمالي الذي يستخدم لرفض الفرضية الأساسية الصحيحة. بمعنى آخر، تمثل قيم مستوى الدلالة (∝) احتمال الوقوع في النوع الأول من الخطأ (Type 1 Error). وعليه، فقد تمّ تحديد قيمة ∝ بالمثال بـ ٥٪ . يجب ملاحظة هنا، انه كلما صغرت قيمة ∝ كلما زادت قيمة ß لأي حجم عينة.

الخطوة الرابعة : ايجاد القيمة المعيارية للاختبار الاحصائي.

ان القيم المعيارية للاختبار الاحصائي (Z) موجودة في الجدول رقم (٢) للملحق الموجود في نهاية الكتاب للاختبار ذو الطرفين والجدول رقم (٣) في نفس الملحق للاختبار ذو الطرف الواحد. ان هذه القيم الموجودة في هذه الجداول معطاة للوسط صفر والانحراف المعياري واحد. لذلك، فان استخدام هذه الجداول يتطلب تحويل البيانات من العينة لازالة تاثير وحدة القياس. ان تقاطع القيم الموجودة في العمود الأول والصف الأول يحدد قيمة أو قيم (Z) (بمعنى آخر، عدد الانحرافات المعيارية عن الوسط ذات الاهتمام بالنسبة للشخص الباحث). فعلى سبيل المثال، نجد ان قيمة Z المعيارية اذا كانت قيمة ∝ = ٥٪ عن طريق استخدام الجدول رقم (٢) هي (١ر٩٦) (١ر٩ من العمود الأول و ٠٦ر من الصف الأول) انحرافات معيارية عن الوسط. اننا نلاحظ هنا ان هذا يمثل مستوى ثقة ٩٥٪. انه لمن الجدير بالذكر أن نذكر هنا ان فترة الثقة ٩٥٪ هي (سَ ± ٢ ضرب الخطأ المعياري، وفي الحقيقة ان فترة الثقة هي سَ ± ١ر٩٦ ضرب الخطأ المعياري) .

ان ما تعنيه القيمة ١ر٩٦ هو ان احتمال الحصول على قيمة لـ(Z) أكبر من ١ر٩٦ هو أقل من ٥٪ كما اننا نلاحظ ان احتمال الحصول على قيمة لـ(Z) أكبر من ١ر٩٦ من الجدول رقم (٣) هو أقل من (٠٢٥ر). لهذا، فانه يمكن استخدام الاختبار هنا عندما تكون قيمة ∝ = (٠٢٥ر). أما اذا كان الاختبار ذو طرف

واحد، فاننا نلاحظ ان قيمة Z = (١٫٦٤) اذا كانت قيمة ∝ = ٥٪. لهذا، فان احتمال الحصول على قيمة لـZ أكبر من (١٫٦٤) لاختبار ذو طرف هو أقل من ٥٪ .

الخطوة الخامسة : ايجاد القيمة الاحصائية للاختبار الاحصائي.

اذا كان الانحراف المعياري لمجتمع الدراسة معروف، فان :

$$Z = \frac{\text{سَ} - \mu}{\sigma_{\bar{x}}}$$

$$Z = \frac{\text{سَ} - \mu}{\sigma / \sqrt{\text{ن}}}$$

اما اذا كان الانحراف المعياري لمجتمع الدراسة مجهولاً، فان :

$$Z = \frac{\text{سَ} - \mu}{s_{\bar{x}}}$$

$$Z = \frac{\text{سَ} - \mu}{s / \sqrt{\text{ن}}}$$

ان قيمة الانحراف المعياري لمجتمع الدراسة (σ) في مثالنا هذه هي قيمة مجهولة، لهذا فاننا سوف نستخدم المعادلة الثانية. ان ما تقوم به المعادلة هو عرض الفرق ما بين القيمة المحسوبة للوسط الحسابي (سَ) والقيمة االمفترضة للوسط (μ) كمقياس يعبر عنه بعدد الأخطاء المعيارية. ان السؤال الذي يطرح نفسه هنا هو : هل الفرق ما بين سَ – μ والمعبر عنه بالخطأ المعياري (قيمةZالمحسوبة) كبير وهل السبب وفي حدوثه هو الخطأ العيني والذي احتماله هو أقل من ٥٪ ؟

$$Z = \frac{٢٤ - ٢٣}{٥ / \sqrt{١٠٠}}$$

$$= \frac{١}{٥ / ١٠} = ٢ر$$

لهذا، فان الفرق ما بين ٢٤ و ٢٣ هو خطاين معيارين.

الخطوة السادسة : مقارنة قيم (Z)

عند مقارنة قيم (Z) المحسوبة والمعيارية، نلاحظ ان القيمة المحسوبة لـ Z أكبر من القيمة المعيارية لنفس الاختبار عندما كانت قيمة α = ٥٪ (٢ > ٩٦ر١) لذلك، فان القرار سوف يكون رفض الفرضية الأساسية وقبول الفرضية البديلة. والمعنى الاداري لهذا القرار هو عدم إمكانية اعتبار العمر (٢٣) سنة كمعدّل عمري لمجتمع الدراسة.

ـ اختبار t

ان استخدام اختبار t يعتبر مناسباً للاختبار الفرضيات المتعلّقة بالأوساط بالنسبة لجميع احجام العينات عندما يكون الانحراف المعياري بالنسبة لمجتمع الدراسة مجهول. ان السبب في استخدام اختبار Z بدلاً من اختبار t عندما يكون حجم العينة أكبر من (٣٠) هو ان توزيع t وتوزيع Z هما توزيعان متطابقان عندما يكون حجم العينة أكبر من (٣٠)، كما ان قيم t بالنسبة لتوزيع t لم يتم حسابها بالنسبة للعينات كبيرة الحجم. ففي اختبار t نقوم بتقدير قيمة الانحراف المعياري لمجتمع الدراسة (σ) بقيمة الانحراف المعياري لعينة الدراسة (S). ان القيم المعيارية للاختبار الاحصائي t مزودة بالجدول رقم (٤) الموجود في الملحق. ان قيمة t المعيارية تتغير مع تغيّر مستوى الثقة (α) الذي يتم اختياره، وكذلك تتغيّر مع تغير درجات الحرية بالنسبة للعينة، هذا بالاضافة الى تاثر القيمة بنوع الاختبار المطلوب (اختبار t ذو طرف أم اختبار ذو

طرفين). فعلى سبيل المثال، ان قيمة t المعيارية لاختبار ذو طرف واحد اذا كانت قيمة ∝ = ٥٪ ودرجات الحرية تساوي (١٠) هي t = ١٫٨١٢ . ان درجات الحرية للاختبار t بالنسبة لاختبار الوسط هي دائماً (ن – ١) والسبب في ذلك استخدام الانحراف المعياري للعينة (S) لتقدير الانحراف المعياري لمجتمع الدراسة (σ)، ودرجات الحرية لـ (S) هي (ن – ١). وسوف نقوم بتوضيح اختبار t من خلال المثال التالي :

يريد طالب عمل دراسة لمعرفة متوسط استهلاك الفرد السنوي من الماء بالجالون. فاختار عينة عشوائية مكونة من (٧) أفراد، ووجد أن متوسط الاستهلاك السنوي من الماء بالجالون كان (١٢٠) جالون والانحراف المعياري كان (١٥). فهل يستطيع الطلب اعهتبار متوسط الاستهلاك السنوي هو (١٠٠) جالون بالاعتماد على متوسط استهلاك العينة، عند مستوى الدلالة (∝) ٥٪ .

الحل : سَ = ١٢٠ ، S = ١٥ ، ن = ٧ ، ∝ = ٥٪

١ ـ تشكيل الفرضية الأساسية والفرضية البديلة

ف. : μ = ١٠٠

ف١ : μ ≠ ١٠٠

٢ ـ تحديد مستوى الدلالة ∝

∝ = ٥٪

٣ ـ تحديد الاختبار الأمثل.

اختبار t ذو الطرفين هو الاختبار الأمثل، لان الانحراف المعياري لمجتمع الدراسة مجهول وحجم العينة أقل من (٣٠) .

٤ ـ ايجاد القيمة المعيارية لاختبار t .

حتى نتمكن من ايجاد القيمة المعيارية، فانه لا بد من معرفة درجات الحرية ومستوى الدلالة ($\propto$) .

درجات الحرية (د . ح) = ن – ١

= ٧ – ١ = ٦

مستوى الدلالة = ٥٪

t (د . ح = ٦، $\propto$ = ٥٪) = ٢٫٤٤٧ (من الجدول رقم ٤) .

٥ ـ ايجاد القيمة الاحصائية للاختبار t

$$t = \frac{\bar{س} - \mu}{s_{\bar{x}}}$$

$$t = \frac{\bar{س} - \mu}{s / \sqrt{ن}}$$

$$t = \frac{١٢٠ - ١٠٠}{١٥ / \sqrt{٧}}$$

$$t = ٣٫٥٣$$

٦ ـ القرار : بما أن القيمة المحسوبة للاختبار أكبر من القيمة المعيارية، فان هذا يعني رفض الفرضية الأساسية وقبول الفرضية البديلة . وعليه، فانه لا يمكن اعتبار متوسط استهلاك الفرد السنوي من الماء (١٠٠) جالون.

مثال : افرض انه في المثال السابق كان المطلوب هو معرفة ما اذا كان بالامكان اعتبار ان

متوسط حجم الاستهلاك السنوي للفرد من الماء أقل من (١٠٠) جالون، بالاعتماد على متوسط استهلاك العينة.

وعند نفس مستوى الدلالة ∝ ٥٪. فهل يستطيع الطالب عمل ذلك.

الحل : ١) ف. : $\mu \geq$ ١٠٠

ف$_١$: $\mu <$ ١٠٠

٢) ∝ = ٥٪

٣) الاختبار الأمثل هو اختبار t ذو الطرف الواحد.

٤) القيمة المعيارية لاختبار t هي :

t (د . ح = ٦ ، ∝ = ٥٪) = ١٫٩٤٣ (من الجدول رقم ٤) .

٥) القيمة الاحصائية لاختبار t = ٣٫٥٣

٦) القرار : رفض الفرضية الأساسية وقبول الفرضية البديلة لان القيمة الاحصائية أكبر من القيمة المعيارية. وعليه، فانه لا يمكن اعتبار ان متوسط الاستهلاك السنوي للفرد من الماء أقل من (١٠٠) جالون.

ـ اختبار الأوساط لعينتين كبيرتا الحجم : Two-Sample Test of Means : Large Samples

لقد قمنا بالسابق بوصف كيفية تشكيل الاجراء الاحصائي لاختبار ما اذا كان الوسط الحسابي لمجتمع الدراسي مساوياً لقيمة محددة Specified Value . أما الآن، فاننا سوف نقوم بالتطرق الى الحالة التي تكون فيها العينة العشوائية قم تمّ اختيارها من مجتمعين دراسين. لهذا، فاننا سوف نهتم في هذا الجزء بوصف كيفية اختبار الفرضية التي تعتبر ان الأوساط الحسابية لكلا المجتمعين متساويان. كما اننا سوف نفترض هنا ان حجم كل عينة كبير (ن > ٣٠). وسوف نقوم بتوضيح هذا النوع من الاختبار من

خلال المثال التالي :

مثال : شركة لها منشأتان، وكل منشأة من هاتين المنشأتين تقوم بانتاج الصحون. لذلك. فانه من الضروري للشركة اختبار ما اذا كان متوسط قطر الصحون المنتجة في المنشأة الأولى مساوياً لمتوسط قطر الصحون المنتجة بالمنشأة الثانية. فاذا كان متوسط قطر الصحون في أي منشأة أكبر من متوسط قطر الصحون في المنشأة الأخرى، فان هذا يعني رفض الفرضية التي تعتبر ان متوسط الأقطار للصحون في كلا المنشأتين متساوي. افرض ان الانحراف المعياري في أقطار الصحون المنتجة لكلا المنشأتين هو ٢٥ سم. وقد تم اختيار عينة عشوائية من منتجات كل منشأة حجمها (١٠٠) صحن، ووجد ان متوسط قطر الصحون لإحدى المنشأتين كان (٤٠) سم والأخرى (٤٥) سم. هل ترفض الشركة الفرضية التي تعتبر ان الأوساط متساوية، عند مستوى الدلالة ($\propto$) ٥٪ ؟

الحل : افرض ان :

μ_1 : متوسط قطر الصحون المنتجة في المنشأة الأولى.

μ_2 : متوسط قطر الصحون المنتجة في المنشأة الثانية.

١ - تشكيل الفرضية الأساسية والبديلة.

$$ف_0 : \mu_1 = \mu_2 \quad \text{أو} \quad \mu_1 - \mu_2 = \text{صفر}$$

$$ف_1 : \mu_1 \neq \mu_2 \quad \text{أو} \quad \mu_1 - \mu_2 \neq \text{صفر}$$

٢ - اختيار الاختبار الأمثل.

الاختبار الأمثل هو اختبار Z ذو الطرفين.

اذا كان حجم كلا العينتين كبير، فقد ثبت ان التوزيع العيني للفرق ما بين الأوساط العينية ($\bar{س}_1 - \bar{س}_2$) هو تقريباً توزيع طبيعي وبوسط حسابي هو ($\mu_1 - \mu_2$) وانحراف معياري $\sqrt{\sigma_1^2 / ن_1 + \sigma_2^2 / ن_2}$. لهذا، اذا كانت

الفرضية الأساسية صحيحة (والتي تعني ان $\mu_1 - \mu_2$ = صفر)، فان الفرق ما بين الأوساط الحسابية يكون تقريباً موزع طبيعياً وبوسط حسابي مساوياً للصفر وانحراف معياري قدره $\sqrt{\sigma_1^2 / ن_1 + \sigma_2^2 / ن_2}$.

٣ ـ تحديد مستوى الدلالة ($\propto$) .

لقد تمّ تحديد قيمة $\propto$ بالمثال عند مستوى ٥٪ .

٤ ـ ايجاد القيمة المعيارية للاختبار الاحصائي.

باستخدام الجدول رقم (٢) والموجود في الملحق، نجد ان قيمة Z المعيارية عند مستوى الدلالة ٥٪ للاختبار ذو الطرفين هي (١ر٩٦) .

٥ ـ ايجاد القيمة الاحصائية للاختبار الاحصائي.

يمكن ايجاد القيمة الاحصائية للاختبار اذا كانت العينتين كبيرتا الحجم من خلال العلاقة الاحصائية التالية :

$$Z = (\bar{س}_1 - \bar{س}_2) \div \sqrt{\sigma_1^2 / ن_1 + \sigma_2^2 / ن_2}$$

$$= (٤٠ - ٤٥) \div \sqrt{٥^2 / ١٠٠ + ٥^2 / ١٠٠}$$

$$= -٥ \div \sqrt{\frac{١}{٤} + \frac{١}{٤}}$$

$$= -٥ \div \sqrt{\frac{١}{٢}}$$

$$= -٤٠ر٧$$

٦ ـ القرار : بما أن القيمة المطلقة للقيمة الاحصائية أكبر من القيمة المعيارية للاختبار، فان هذا يعني رفض الفرضية الأساسية وقبول الفرضية البديلة. وبمعنى آخر،فانه لا يمكن اعتبار متوسط أقطار الصحون المنتجة في كلا المنشآتين متساوي.

ـ اختبار الأوساط لعينتين صغيرتا الحجم : **Two-Sample Test of Means : Small Samples**

سوف نقوم الآن بوصف كيفية اختبار الفرضيات التي تعتبر ان الأوساط الحسابية للمجتمعات الدراسية متساوية اذا كانت العينات صغيرة الحجم والانحراف المعياري لمجتمعات الدراسة غير معروف. ان الاختبارات تفترض ان الانحرافات المعيارية للمجتمعين متساوية. وسوف نقوم بتوضيح هذا النوع من الاختبارات من خلال المثال التالي :

مثال : تريد إحدى شركات السيارات اختبار نوعين من المنتجات التي تنتجها ولتكن نوع أ ونوع ب، لتحديد ما اذا كان متوسط السرعة لكلا النوعين تحت مجموعة معينة من الظروف هو نفسه. وعليه، فقد تمّ اجراء أربعة اختبارات للسرعة بالنسبة للنوع أ تحت نفس الظروف التي تمّ تحديدها ووجد ان متوسط السرعة كانت ١٨٠ كم /بالساعة والانحراف المعياري ٣٠ كم /بالساعة. كما وقد تم اجراء أربعة اختبارات للسرعة بالنسبة للنوع ب وتحت نفس الظروف ووجد ان متوسط السرعة ٢١٠ كم /بالساعة والانحراف المعياري ٢٠ كم /بالساعة. فهل تستطيع الشركة قبول الفرضية التي تعتبر ان متوسط السرعة لكلا النوعين متساوي، اذا كانت قيمة α = ٥٪ ؟

الحل : افرض ان :

μ_1 : متوسط السرعة للنوع أ من السيارات.

μ_2 : متوسط السرعة للنوع ب من السيارات.

١ ـ تشكيل الفرضية الأساسية والبديلة.

$\text{ف}_{.} : \mu_{١} = \mu_{٢}$ أو $\mu_{١} - \mu_{٢} =$ صفر.

$\text{ف}_{١} : \mu_{١} \neq \mu_{٢}$ أو $\mu_{١} - \mu_{٢} \neq$ صفر

٢ ـ اختيار الاختبار الأمثل.

الاختبار الأمثل هو اختبار t ذو الطرفين.

اذا كان المجتمعان موزعان توزيعاً طبيعياً والانحراف المعياري لها متساوي، وكانت الفرضية الأساسية صحيحة، فقد ثبت أن

$(\bar{\text{س}}_{١} - \bar{\text{س}}_{٢}) \div \sqrt{S^{٢} (١/\text{ن}_{١} + ١/\text{ن}_{٢})}$ لها توزيع (t) وبدرجات حرية = $\text{ن}_{١} + \text{ن}_{٢} - ٢$ علماً بأن :

$$S^{٢} = \frac{(\text{ن}_{١} - ١)(S_{١}^{٢}) + (\text{ن}_{٢} - ١)(S_{٢}^{٢})}{\text{ن}_{١} + \text{ن}_{٢} - ٢}$$

لذلك، لاختبار الفرضية الأساسية، فانه لا بد من حساب

$$\tau = (\bar{\text{س}}_{١} - \bar{\text{س}}_{٢}) \div \sqrt{S^{٢} (١/\text{ن}_{١} + ١/\text{ن}_{٢})}$$

٣ ـ تحديد مستوى الدلالة ($\propto$)

لقد تمّ تحديد مستوى الدلالة ($\propto$) بالمثال بـ ٥٪ .

٤ ـ ايجاد القيمة المعيارية للاختبار الاحصائي.

لايجاد القيمة المعيارية للاختبار الاحصائي (t) ذو الطرفين فانه لا بد من معرفة كل من :

أ ـ درجات الحرية = ن$_1$ + ن$_2$ − ٢

= ٤ + ٤ − ٢

= ٦

ب ـ مستوى الدلالة (α) = ٥٪

∴ t (د . ح = ٦ ، α = ٥٪) = ٢٫٤٤٧

٥ ـ ايجاد القيمة الاحصائية للاختبار

يمكن ايجاد القيمة الاحصائية للاختبار اذا كانت العينتين صغيرتا الحجم والمجتمعين موزعين توزيعاً طبيعياً والانحراف المعياري لهما متساوي من خلال العلاقة الاحصائية التالية :

$$t = (\bar{س}_1 - \bar{س}_2) \div \sqrt{S^2 (١ / ن_1 + ١ / ن_2)}$$

ولايجاد قيمة t ، لا بد من ايجاد قيمة S^2 أولاً من خلال العلاقة التالية :

$$S^2 = \frac{(ن_1 - ١)(S_1^2) + (ن_2 - ١)(S_2^2)}{ن_1 + ن_2 - ٢}$$

$$S^2 = \frac{(٤ - ١)(٣٠)^2 + (٤ - ١)(٢٠)^2}{٤ + ٤ - ٢}$$

$$S^2 = \frac{٢٧٠٠ + ١٢٠٠}{٦}$$

$$S^2 = ٦٥٠$$

إذن $T = (١٨٠ - ٢١٠) \div \sqrt{٦٥٠ (\frac{١}{٤} + \frac{١}{٤})}$

$= (-٣٠) \div \sqrt{٣٢٥}$

$= (-٣٠) \div ١٨٫٠٣$

$= -١٫٦٦$

٦ ـ القرار : بما أن القيمة الاحصائية المطلقة للاختبار الاحصائي أقل من القيمة المعيارية، فهذا يعني قبول الفرضية الأساسية ورفض البديلة . بمعنى آخر، هناك امكانية لاعتبار أن متوسط السرعة لكلا النوعين من السيارات متساوي .

ـ اختبار كاي تربيع Chi-Square Test

ان أحد التطبيقات المهمة لتوزيع كاي تربيع ($X^{٢}$) هو في المسائل التي يكون فيها الشخص متخذ القرار مهتم بتحديد ما اذا كانت النسب المختلفة متساوية . لذلك . فان الاهتمام هنا سوف يتحول من التركيز على اختبار ما اذا كانت قيمتين لمجتمعين مختلفتين متساوية الى التركيز على كيفية اختبار ما اذا كانت أكثر من قيمتين أو نسبتين متساوية .

وعليه، فان شكل الفرضية الأساسية في هذا النوع من المسائل يكون كما يلي :

ف. : $\pi_{١} = \pi_{٢} = \pi_{٣} \dots = \pi_{ر}$

أما الفرضية البديلة فانها تعتبر أن هذه النسب غير متساوية وكما يلي :

ف١ : $\pi_{١} \neq \pi_{٢} \neq \dots \neq \pi_{ر}$

ولتوضيح هذا النوع من الاختبارات، دعنا نقوم بحل المثال التالي :

تقوم إحدى الشركات باختيار عينة عشوائية مكونة من (٥٠) وحدة من

منتجات إحدى السلع لمعرفة عدد الوحدات المعيبة في الانتاج اليومي. فاذا تمّ اختيار عينة لمدة (٢٨) يوماً وكانت الوحدات المعيبة في كل عينة مزودة بالجدول (١١ - ٢). فالسؤال الذي يطرح نفسه الآن وتريد ادارة الشركة اجابة عليه هو : هل نسبة الوحدات المعيبة لجميع الأيام ثابتة؟ فاذا تبين ان النسبة تتغير من يوم الى آخر، فان هذا يعني ان العملية الانتاجية ليست تحت السيطرة. وعليه، فانه لا بد للشركة من القيام بالمحاولة لمعرفة الأسباب التي تؤدي الى التغيّر.

ان اهتمام الفرضية الأساسية بمثل هذا النوع من المشاكل يكون حول مساواة جميع النسب. بمعنى آخر، يكون شكل الفرضية الأساسية كما يلي :

$$ف_{٠} : \pi_{١} = \pi_{٢} = \ldots = \pi_{ر}$$

بينما الفرضية البديلة تقول بان هذه النسب ليست متساوية ،كما يلي :

$$ف_{١} : \pi_{١} \neq \pi_{٢} \neq \ldots \neq \pi_{ر}$$

حيث ان :

$\pi_{١}$ تمثل نسبة المعيب في الانتاج الكلي لليوم الأول

$\pi_{٢}$ تمثل نسبة المعيب في الانتاج الكلي لليوم الثاني، وهكذا.

فاذا كانت الفرضية الأساسية صحيحة، فان نسبة المعيب لجميع الأيام يمكن تقديره من الوسط الحسابي للوحدات المعيبة لجميع الأيام وباستخدام العلاقة التالية :

$$\left(\sum س_{ر} \right) / \sum ن_{ر}$$

حيث ان :

$س_{ر}$: عدد الوحدات المعيبة في اليوم (ر) .

$ن_{ر}$: حجم العينة في اليوم (ر) .

جدول (١١ - ٢)

عدد الوحدات المعيبة في (٢٨) يوم عمل

اليوم	عدد الوحدات المعيبة (ف)	عدد الوحدات السليمة (ف)	حجم العينة
١	٤	٤٦	٥٠
٢	٩	٤١	٥٠
٣	١٠	٤٠	٥٠
٤	١١	٣٩	٥٠
٥	١٣	٣٧	٥٠
٦	٣٠	٢٠	٥٠
٧	٢٦	٢٤	٥٠
٨	١٣	٣٧	٥٠
٩	٨	٤٢	٥٠
١٠	٢٣	٢٧	٥٠
١١	٣٤	١٦	٥٠
١٢	٢٥	٢٥	٥٠
١٣	١٨	٣٢	٥٠
١٤	١٢	٣٨	٥٠
١٥	٤	٤٦	٥٠
١٦	٣	٤٧	٥٠
١٧	١١	٣٩	٥٠
١٨	٨	٤٢	٥٠
١٩	١٤	٣٦	٥٠
٢٠	٢١	٢٩	٥٠
٢١	٢٥	٢٥	٥٠
٢٢	١٨	٣٢	٥٠
٢٣	١٠	٤٠	٥٠
٢٤	٨	٤٢	٥٠
٢٥	١٨	٣٢	٥٠
٢٦	١٩	٣١	٥٠
٢٧	٤	٤٦	٥٠
٢٨	٨	٤٢	٥٠

وفي مثالنا هذا، فان نسبة الوحدات المعيبة المتوقعة لأي يوم من الأيام، اذا كانت الفرضية الأساسية صحيحة، هي :

$$\frac{٤ + ٩ + ١٠ + ١١ + \ldots + ١٩ + ٤ + ٨}{٥٠ + ٥٠ + ٥٠ + \ldots + ٥٠ + ٥٠ + ٥٠} = ٢٩و٠$$

لاختبار ما اذا كانت الفرضية الأساسية صحيحة، فاننا نقوم بحساب عدد الوحدات المعيب المتوقع وغير المعيب المتوقع لكل يوم. بما ان نسبة المعيب المتوقعة هي (٢٩ر)، فاننا نتوقع ان يكون عدد الوحدات المعيبة في كل يوم هي (٥٠) (٢٩ر) = (٥ر١٤). وعدد الوحدات السليمة المنتجة في كل يوم هي (٥٠) (٧١ر) = (٥ر٣٥). ولاختبار الفرضية الأساسية، فان هذا يتطلب مقارنة النسبة المفترضة مع النسبة الحقيقية. وبوضوح، كلما زاد الفرق ما بين النسبة المفترضة والنسبة الحقيقية، كلما قل احتمال ان تكون الفرضية الأساسية صحيحة.

وبعد حساب تكرار الوحدات المعيبة المتوقع، فانه لا بد من حساب القيمة الاحصائية التالية وكما يظهر في الجدول (١١ -٣) :

$$\sum (ف - و)^{٢} / و$$

حيث ان :

ف : تكرار الوحدات المعيبة الحقيقي.

و : التكرار المتوقع للوحدات المعيبة

فاذا كانت الفرضية الأساسية صحيحة، فان التوزيع العيني للاختبار الاحصائي يمكن تقريبه بتوزيع كاي تربيع وبدرجات حرّية تساوي (ر - ١)، حيث ان ر تمثل عدد نسب مجتمع الدراسة التي تمّ مقارنتها.

ويتم رفض الفرضية الأساسية التي تعتبر ان $\pi_{١} = \pi_{٢} = \ldots = \pi_{ر}$ اذا كان

$\sum$ (ف – و)^{2} > X^{2}_{α} ، حيث ان ∞ تمثل مستوى الدلالة للاختبار وبدرجات تساوي (ر – ١). ويتم قبول الفرضية البديلة اذا كان $\sum$ (ف – و)^{2} > X^{2}_{α} .

يبين الجدول رقم (١١ - ٣) ان قيمة $\sum$ (ف – و)^{2} / و هي تقريباً ١٨٢ .

جدول رقم (١١ - ٣)

النسب المتوقعة وحساب $\sum$ (ف – و)^{2} / و

اليوم	عدد الوحدات المعيب			عدد الوحدات السليمة		
	الفعلي (ف)	المتوقع (و)	(ف – و)^{2} / و	الفعلي (ف)	المتوقع (و)	(ف – و)^{2} / و
١	٤	١٤٫٥	(٤ – ١٤٫٥)^{2}/ ١٤٫٥	٤٦	٣٥٫٥	(٤٦ – ٣٥٫٥)^{2}/ ٣٥٫٥
٢	٩	١٤٫٥	(٩ – ١٤٫٥)^{2}/ ١٤٫٥	٤١	٣٥٫٥	(٤١ – ٣٥٫٥)^{2}/ ٣٥٫٥
٣	١٠	١٤٫٥	(١٠ – ١٤٫٥)^{2}/ ١٤٫٥	٤٠	٣٥٫٥	(٤٠ – ٣٥٫٥)^{2}/ ٣٥٫٥
٤	١١	١٤٫٥	(١١ – ١٤٫٥)^{2}/ ١٤٫٥	٣٩	٣٥٫٥	(٣٩ – ٣٥٫٥)^{2}/ ٣٥٫٥
٥	١٣	١٤٫٥	(١٣ – ١٤٫٥)^{2}/ ١٤٫٥	٣٧	٣٥٫٥	(٣٧ – ٣٥٫٥)^{2}/ ٣٥٫٥
٦	٣٠	١٤٫٥	(٣٠ – ١٤٫٥)^{2}/ ١٤٫٥	٢٠	٣٥٫٥	(٢٠ – ٣٥٫٥)^{2}/ ٣٥٫٥
٧	٢٦	١٤٫٥	(٢٦ – ١٤٫٥)^{2}/ ١٤٫٥	٢٤	٣٥٫٥	(٢٤ – ٣٥٫٥)^{2}/ ٣٥٫٥
٨	١٣	١٤٫٥	(١٣ – ١٤٫٥)^{2}/ ١٤٫٥	٣٧	٣٥٫٥	(٣٧ – ٣٥٫٥)^{2}/ ٣٥٫٥
٩	٨	١٤٫٥	(٨ – ١٤٫٥)^{2}/ ١٤٫٥	٤٢	٣٥٫٥	(٤٢ – ٣٥٫٥)^{2}/ ٣٥٫٥
١٠	٢٣	١٤٫٥	(٢٣ – ١٤٫٥)^{2}/ ١٤٫٥	٢٧	٣٥٫٥	(٢٧ – ٣٥٫٥)^{2}/ ٣٥٫٥
١١	٣٤	١٤٫٥	(٣٤ – ١٤٫٥)^{2}/ ١٤٫٥	١٦	٣٥٫٥	(١٦ – ٣٥٫٥)^{2}/ ٣٥٫٥
١٢	٢٥	١٤٫٥	(٢٥ – ١٤٫٥)^{2}/ ١٤٫٥	٢٥	٣٥٫٥	(٢٥ – ٣٥٫٥)^{2}/ ٣٥٫٥

تابع / جدول رقم (١١ - ٣)

اليوم	عدد الوحدات المعيب			عدد الوحدات السليمة		
	الفعلي (ف)	المتوقع (و)	(ف - و)^٢ / و	الفعلي (ف)	المتوقع (و)	(ف - و)^٢ / و
١٣	١٨	١٤،٥	(١٨ - ١٤،٥)^٢ / ١٤،٥	٣٢	٣٥،٥	(٣٢ - ٣٥،٥)^٢ / ٣٥،٥
١٤	١٢	١٤،٥	(١٢ - ١٤،٥)^٢ / ١٤،٥	٣٨	٣٥،٥	(٣٨ - ٣٥،٥)^٢ / ٣٥،٥
١٥	٤	١٤،٥	(٤ - ١٤،٥)^٢ / ١٤،٥	٤٦	٣٥،٥	(٤٦ - ٣٥،٥)^٢ / ٣٥،٥
١٦	٣	١٤،٥	(٣ - ١٤،٥)^٢ / ١٤،٥	٤٧	٣٥،٥	(٤٧ - ٣٥،٥)^٢ / ٣٥،٥
١٧	١١	١٤،٥	(١١ - ١٤،٥)^٢ / ١٤،٥	٣٩	٣٥،٥	(٣٩ - ٣٥،٥)^٢ / ٣٥،٥
١٨	٨	١٤،٥	(٨ - ١٤،٥)^٢ / ١٤،٥	٤٢	٣٥،٥	(٤٢ - ٣٥،٥)^٢ / ٣٥،٥
١٩	١٤	١٤،٥	(١٤ - ١٤،٥)^٢ / ١٤،٥	٣٦	٣٥،٥	(٣٦ - ٣٥،٥)^٢ / ٣٥،٥
٢٠	٢١	١٤،٥	(٢١ - ١٤،٥)^٢ / ١٤،٥	٢٩	٣٥،٥	(٢٩ - ٣٥،٥)^٢ / ٣٥،٥
٢١	٢٥	١٤،٥	(٢٥ - ١٤،٥)^٢ / ١٤،٥	٢٥	٣٥،٥	(٢٥ - ٣٥،٥)^٢ / ٣٥،٥
٢٢	١٨	١٤،٥	(١٨ - ١٤،٥)^٢ / ١٤،٥	٣٢	٣٥،٥	(٣٢ - ٣٥،٥)^٢ / ٣٥،٥
٢٣	١٠	١٤،٥	(١٠ - ١٤،٥)^٢ / ١٤،٥	٤٠	٣٥،٥	(٤٠ - ٣٥،٥)^٢ / ٣٥،٥
٢٤	٨	١٤،٥	(٨ - ١٤،٥)^٢ / ١٤،٥	٤٢	٣٥،٥	(٤٢ - ٣٥،٥)^٢ / ٣٥،٥
٢٥	١٨	١٤،٥	(١٨ - ١٤،٥)^٢ / ١٤،٥	٣٢	٣٥،٥	(٣٢ - ٣٥،٥)^٢ / ٣٥،٥
٢٦	١٩	١٤،٥	(١٩ - ١٤،٥)^٢ / ١٤،٥	٣١	٣٥،٥	(٣١ - ٣٥،٥)^٢ / ٣٥،٥
٢٧	٤	١٤،٥	(٤ - ١٤،٥)^٢ / ١٤،٥	٤٦	٣٥،٥	(٤٦ - ٣٥،٥)^٢ / ٣٥،٥
٢٨	٨	١٤،٥	(٨ - ١٤،٥)^٢ / ١٤،٥	٤٢	٣٥،٥	(٤٢ - ٣٥،٥)^٢ / ٣٥،٥
Σ (ف - و)^٢ / و = ١٨٢						

وبالنظر الى الجدول رقم (٦) والذي يبين القيم المعيارية للاختبار كاي تربيع، نجد أن قيمة X^2 اذا كانت قينة ∝ = ٥٪ ودرجات الحرية هي ٢٧ تساوي (١ر٤٠). وبما ان القيمة المحسوبة (١٨٢) أكبر من القيمة المعيارية (١ر٤٠)، فانه من الواضح جداً ان نقوم برفض الفرضية الأساسية. وهذا يعني ان نسب الوحدات المعيبة في الانتاج تختلف من يوم الى خر، والذي يعني ان انجاز هذه الشركة ليس تحت السيطرة أو المراقبة. وبناءاً على هذه النتائج، فان هذا سوف يؤدي بالشركة لعمل دراسة لمعرفة الأسباب التي تؤدي الى التباين في الإنجاز لانه تبين ان نسب الاختلاف بالانتاج لم يكن بمحض الصدفة.

مثال : يريد باحث اقتصادي اختبار الفرضية التي تعتبر أن نسبة المنشآت التي تعمد الى رفع الأسعار في السنة القادمة لثلاثة صناعات أ ، ب ، جـ هي نفس النسبة . فاذا كانت البيانات المتوفرة عن عينات من المنشآت هي كما يلي :

القرار	عدد المنشآت		
	الصناعة أ	الصناعة ب	الصناعة جـ
مع رفع الأسعار	٤٠ منشأة	٥٠ منشأة	٦٠ منشأة
مع عدم رفع الأسعار	٦٠ منشأة	٥٠ منشأة	٤٠ منشأة

فهل يستطيع الاقتصادي قبول أم رفض هذه الفرضية، اذا علمت ان مستوى الدلالة هو ٥٪ .

الحل : ان عدد المنشآت التي مع رفع السعر تساوي ١٥٠ / ٣٠٠. لهذا، فان النسبة الكلية للمنشآت التي مع رفع السعر هي (٥ر). ان عدد المنشآت التي مع زيادة السعر، اذا كانت الفرضية الأساسية صحيحية، هو

القرار	عدد المنشآت		
	الصناعة أ	الصناعة ب	الصناعة جـ
مع رفع الأسعار	٥٠ منشأة	٥٠ منشأة	٦٠ منشأة
مع عدم رفع الأسعار	٥٠ منشأة	٥٠ منشأة	٤٠ منشأة

لهذا ، $\sum$ (ف – و)^2 / و =

$$= \frac{(40-50)^2}{50} + \frac{(60-50)^2}{50} + \frac{(50-50)^2}{50} +$$

$$= \frac{(50-50)^2}{50} + \frac{(60-50)^2}{50} + \frac{(40-50)^2}{50} =$$

= ٢ + ٢ + ٠ + ٠ + ٢ + ٢

= ٨

بما أن هناك نسب لثلاثة مجتمعات متقارنة معاً، فان عدد درجات الحرية هي (٣ – ١ = ٢). وبناءاً على الجدول رقم (٦) في الملحق، فان X^2 (٠،٠٥) = ٥ر٩٩١ عندما تكون درجات الحرية تساوي ٢. وبما أن القيمة المحسوبة لـ $\sum$(ف – و)$_2$ / و أكبر من القيمة المعيارية للاختبار $X^2_{(5.991)}$ ، فان هذا يعني رفض الفرضية الأساسية. بمعنى آخر، يجب على الاقتصادي رفض الفرضية الأساسية التي تعتبر ان نسبة المنشآت التي تعمد الى رفع الأسعار في السنة القادمة هي نفس النسبة لكل صناعة من الصناعات الثلاث المختلفة.

أسئلة

س١ : ما هو المقصود بالفرضية ؟

س٢ : اذكر الخطوات العامة لاختبار الفرضيات ؟

س٣ : ما هو المقصود بقوّة الاختبار ؟

س٤ : ما هو المقصود بالنوع الأول والنوع الثاني من الخطأ ؟

س٥ : لقد وعَدَ مدير المبيعات موظفي المبيعات بالقيام برحلة خاصة اذا كان معدل البيع اليومي لعامل المبيعات ٥٠٠ دينار أو أكثر. فاختار عينة عشوائية مكونة من (١٠) أشخاص وحصل على النتائج التالية : معدل البيع اليومي لموظفي المبيعات هو (٥٥٠) دينار والانحراف المعياري (١٥٠) دينار. فهل يستطيع مدير المبيعات استنتاج ان جميع موظفي المبيعات قد حققوا الهدف المطلوب، اذا علمت ان قيمة α = ٥٪ .

س٦ : تريد مؤسسة طبية تحديد ما اذا كان أقراص تخفيف الوزن لها تأثير مختلف على الأشخاص اللذين اعمارهم فوق الأربعين سنة عن الأشخاص الذين اعمارهم أقل من أربعين سنة. فقد تمّ اعطاء الأقراص الى (١٢) شخص فوق الأربعين وكان متوسط الخسارة بالوزن ٨ كغم والانحراف المعياري هو ٤ كغم. وكذلك تمّ اعطاء الأقراص الى (١٢) شخص تحت الأربعين وكان متوسط الخسارة هو ١١ كغم والانحراف المعياري هو ٣ كغم. أوجد كما يلي :

أ ـ هل هذه البيانات تنسجم مع الفرضية التي تعتبر ان متوسط التاثير على الوزن هو نفسه لكلا المجموعتين من الأشخاص، اذا كانت قيمة α = ٥٪، والاختبار ذو طرفين ؟

ب ـ هل تتغير الاجابة على المطلوب الأول اذا تغيّر مستوى الثقة من ٥٪ الى ١٪ .

جـ هل تتوقع أن تتغير الاجابة في حالة اختيار عينة من كل فئة قدرها (٤٠) شخص وعند مستوى الدلالة ∝ = ٥٪ ؟

س٧ : أوجد قيمة $X^{٢}_{(٠٫٠٥)}$ عندما تكون درجات الحرّية تساوي ٥ ، ١٠ ، ٢٠ ؟

س٨ : تقوم إحدى شركات السيارات باستيراد العجلات من موردين اثنين، مورد أ ومورد ب . وخلال سنة ١٩٩٤، استلمت شركة السيارات عدد العجلات المعيبة والسليمة من كل مورد والتي كانت كما يلي :

المورد	عدد العجلات السليمة	عدد العجلات المعيبة	عدد العجلات المشتراة
أ	٩٤٠	٨٩	١٫٠٢٩
ب	٧٨٠	٣٢	١٫٢

هل تعتقد بأن الشركة سوف تستلم عجلات معيبة من الموردين أ و ب بنفس الاحتمال؟ افترض ان مستوى الدلالة هو ٥٪ .

الفصل الثاني عشر

الانحدار البسيط والارتباط

Simple Regression and Correlation

الانحدار البسيط والارتباط
Simple Regression and Correlation

ـ مقدمة Introduction

غالباً ما يستخدم الباحثون تحليل الانحدار لتقدير كيفية العلاقة أو التأثير ما بين متغيّر ومتغيّر آخر. بمعنى آخر ان الهدف أو الغرض من استخدام تحليل الانحدار هو لتوضيح العلاقة ما بين المتغيرات المستقلة (Independent Variables) في التحليل والمتغير التابع (Dependent Variable)، ومن ثم القيام بتنبؤ (Prediction) قيم المتغيّر التابع بناءاً على القيم المعروفة بالنسبة للمتغيرات المستقلة. مثال: نتريد مؤسسة القيام بتحديد العلاقة ما بين مبيعاتها (Sales) واجمال الناتج القومي (Gross National Product). أو تريد مثلاً تحديد العلاقة ما بين تكلفة الانتاج الكلية (Total Production Costs) ونسبة المخرجات (Output Rate). ولتقدير هذا النوع من العلاقات فانه لا بد من استخدام تحليل الانحدار. أما اذا أراد الشخص الباحث أن يحدد مدى قوة العلاقة ما بين المتغير (المتغيرات) المستقلة والمتغير التابع، فانه لا بد من استخدام تحليل الارتباط. لذلك يعتبر تحليل الانحدار وتحليل الارتباط من بين أكثر الطرق الاحصائية المستخدمة أهمية واستخداماً.

ـ العلاقة ما بين المتغيرات Relationship Among Variables

ان هناك بعض المسائل المتعلقة بقياسات ومشاهدات لمتغير واحد والبعض الآخر يتعلق بقياسات لمتغيرين أو أكثر. فاذا كان لكل قيمة من قيم المتغير المستقل قيمة مقابلة للمتغير التابع، فان هذه الأزواج المرتبة من القيم تدعى مجتمعاً ذا بعدين ويسمى الزوج

المرتب من القيمة والقيمة المناظرة متغيراً عشوائياً ذا بعدين. ان الأمثلة على المجتمعات كثيرة وخاصة في حقل الادارة والاقتصاد. فمن الأمثلة على ذلك بالاضافة الى ما ورد في المقدمة مثلاً : معرفة العلاقة ما بين الرضا الوظيفي والانتاجية أو معرفة العلاقة ما بين الدخل السنوي والادخار أو معرفة العلاقة ما بين الدعاية وحجم المبيعات.

لذلك يجب على الشخص الباحث في بعض الحالات المهمة الاهتمام بأكثر من متغيّر، وبالتحديد يجب على الباحثين الاهتمام بالعلاقات ما بين المتغيرات. فعلى سبيل المثال : يمكن للشخص الباحث الاهتمام بتحديد ما اذا كان التغيّر في متغير ما، وليكن س، له علاقة بالتغير بمتغير آخر وليكن ص. مثال : هل التغيّر بالزيادة بالنسبة للمتغيّر س يؤثر بالزيادة على المتغير ص؟ ان الطرق المستخدمة في الفصول السابقة تعتبر عديمة الفائدة بالنسبة لمعالجة مسالة مثل هذا النوع. كما أنه من المهم جداً ملاحظة ان العلاقات بين المتغيرات ذات الاهتمام بالنسبة للباحثين نادراً ما تكون علاقات محددة أو ثابتة (Deterministic Relationships). ولتوضيح المقصود بالعلاقات غير المحددة، اعتبر المثال التالي : افرض ان المتغير ص هو المتغير المراد تقديره والمتغير س هو المتغير الذي سوف يُستخدم للقيام بعملية التقدير. فاذا كانت العلاقة ما بين المتغير س والمتغير ص ثابتة، فاننا نطلق على هذا النوع من العلاقات بالعلاقات الثابتة أو المحددة. افرض أن المتغير ص هو محيط المربع والمتغير س هو طول ضلع المربع، فانه يمكن التعبير عن العلاقة ما بين المحيط وطول الضلع كما يلي : ص = ٤ س . ويطلق على هذا النوع من العلاقات بالعلاقاتالمحددة لانه اذا عرفنا طول ضلع المربع فاننا نستطيع تحديد محيطه وبدقة. فاذا كان طول ضلع المربع ٢ سم مثلاً، فان محيطه يساوي ٨ سك لا أكثر ولا أقل.

لذلك غالباً ما يكون اهتمام الباحثين بالعلاقات غير محددة. فاذا كان هناك علاقة ما بين المتغيرات س و ص، فان معدل قيم المتغير ص تميل الى الارتباط مع قيمة المتغيّر س، الا انه من المستحيل التنبؤ بقيمة المتغير ص بدقة بناءاً على قيمة المتغير س. فعلى سبيل المثال : افرض ان المتغير س هو معدّل الدخل السنوي لعائلة والمتغير ص يعني قيمة

الادخار السنوي بالنسبة للعائلة. بشكل عام تزداد قيمة الادخار السنوي للعائلة كلما ازداد دخلها، ويمكن استخدام هذه العلاقة في تقدير حجم الادخار في حالة معرفة الدخل. ولكن هذه العلاقة هي بعيدة كل البعد عن الدقة لأن حجم الادخار السنوي للعائلات ذات الدخل المتساوي مختلف. وهذا يعني صعوبة بل استحالة القيام بتنبؤ قيمة الادخار وبدقة بناءاً على دخل العائلة فقط.

ـ تحليل الانحدار Regression Analysis

يصف تحليل الانحدار الطريقة التي من خلالها يرتبط متغير بمتغيّر آخر، كما أن تحليل الانحدار يستطيع معالجة أكثر من متغيرين وهذا ما سيتم شرحه في الفصل اللاحق، أما في هذا الفصل فسوف يتم التركيز على وصف العلاقة ما بين متغيرين فقط.

ان تحليل الانحدار يقوم على اشتقاق معادلة تسمى بمعادلة خط الانحدار والتي تستخدم في تقدير قيمة المتغير التابع المجهولة بناءاً على القيمة المعلومة للمتغير المستقل. مثال : تريد منشأة انتاج (١٠٠٠) وحدة من منتجاتها خلال الشهر القادم وتريد تقدير مقدار التكاليف المتوقعة. على الرغم من معرفة حجم الانتاج في هذه الحالة الا ان تكاليفها غير معلومة. لذلك يمكن استخدام تحليل الانحدار لتقدير قيمة التكاليف بناءاً على القيمة المعلومة لكمية الانتاج. كما يمكن استخدام تحليل الانحدار لتقدير رأس المال المطلوب لانشاء منشأة بطاقة انتاجية محدّدة. ففي مثالنا الحالي اذا كانت الطاقة الانتاجية للمنشأة معروفة، فانه من الممكن استخدام معادلة خط الانحدار لتقدير مستوى الانفاق المطلوب لتحقيق البرنامج الانتاجي.

ان مصطلح تحليل الانحدار قد جاء من الدراسات التي قام بها الاحصائي الانجليزي فرانسس جالتون (Francis Galton) قبل حوالي ٩٢ سنة من الآن. لقد قام جالتون بدراسة مقارنة أطوال الأباء مع أطوال ذرياتهم. ولقد وجد جالتون بدراسته ان طول أبناء الآباء الطوال جداً كانوا اقصر من ابائهم. بينما طول ابناء الآباء القصار جداً

كانوا أطول من ابائهم. بمعنى آخر ان طول ابناء الأباء الطوال جداً أو القصار جداً كان يتناقص تجاه متوسط الطول لمجتمع الدراسة. وبسبب استخدام جالتون لطول الآباء لتقدير الأبناء، أطلق على هذا النوع من التحليل بتحليل الانحدار.

ـ لوحة الانتشار Scatter Diagram

بما ان تحليل الانحدار يهتم بوصف العلاقة ما بين متغير ومتغير، فانه غالباً ما يبدأ هذا النوع من التحليل بالتعامل مع بيانات متغيرين فقط. افرض ان احد المنشآت تريد تقدير العلاقة ما بين كمية الانتاج الشهري والتكاليف الشهرية. ان أول شيء يجب ان تقوم به المنشأة لعمل ذلك هو تجميع بيانات عن كمية الانتاج والتكاليف لعينة من الأشهر الماضية. افرض ان المنشأة قد قامت بتجميع البيانات المطلوبة لعينة مكونة من تسعة أشهر كما في الجدول (١٢ ـ ١). انه من السهولة والملائمة لتمثيل البيانات من هذا النوع على ما يدعى بلوحة الانتشار (Scatter Diagram). ولرسم لوحة الانتشار فاننا نقوم بتمثيل المتغير المعروف أو المتغير المستقل، والذي هو في هذه الحالة كمية الانتاج، على المحور الأفقي أو السيني. أما بالنسبة للمتغير غير المعروف أو المتغير التابع على المحور العمودي أو المحور الصادي. بالطبع تكون قيم المتغيرات معروفة خلال الفترة التي تمّ تجميع بيانات عنها وهي كمية الانتاج والتكاليف خلال الأشهر التسعة الماضية. ولكن عند استخدام تحليل الانحدار لتقدير العلاقة ما بين الانتاج والتكاليف تكون فقط قيم كمية الانتاج معروفة.

جدول (١٢ - ١)

كمية الانتاج والتكاليف لمنشأة خلال (٩) أشهر

كمية الانتاج (بالطن)	تكاليف الانتاج (آلاف الدنانير)
١	٢
٢	٣
٤	٤
٨	٧
٦	٦
٥	٥
٨	٨
٩	٨
٧	٦

يبين الشكل (١٢ - ١) لوحة الانتشار لكمية الانتاج والتكاليف بناءاً على البيانات الموجودة في الجدول (١٢ - ١). بالنظر الى الشكل (١٢ - ١) نلاحظ ان لوحة الانتشار تزوّد وبوضوح صورة وصفية مفيدة للعلاقة ما بين المتغير التابع والمتغير المستقل. وبالاعتماد على لوحة الانتشار، فان الشخص الباحث يستطيع اشتقاق الانطباع الأولي عن الأسئلة الثلاثة المهمة التالية :

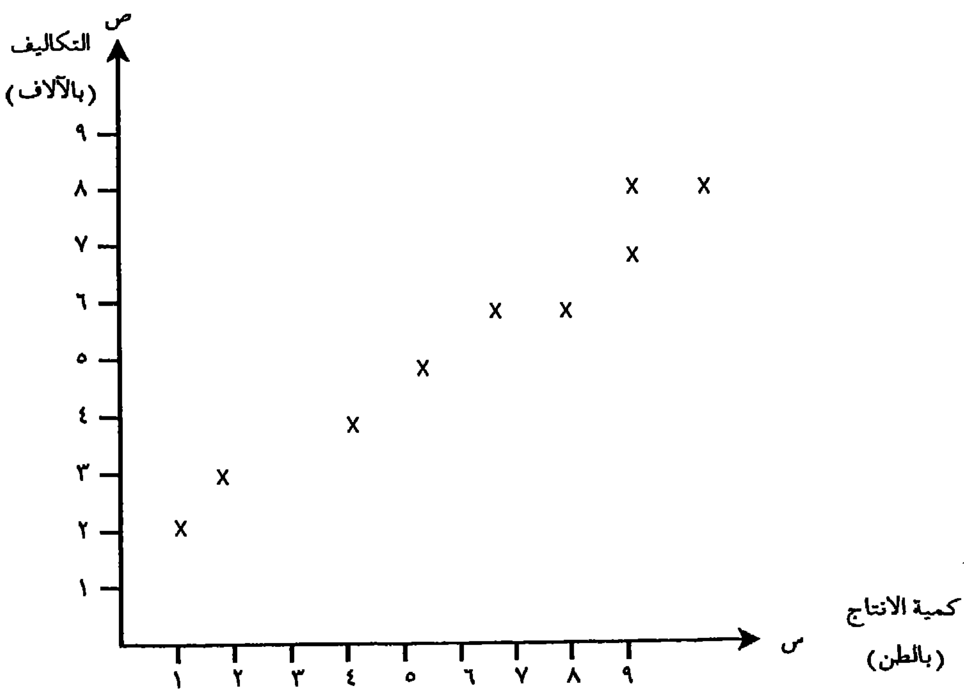
أ - هل العلاقة ما بين المتغيرين طردية أم عكسية ؟

تكون العلاقة ما بين المتغيرين علاقة طردية اذا أدت الزيادة بالمتغير المستقل الى زيادة بالمتغير التابع، أو أدى النقصان في المتغير المستقل الى نقصان بالمتغير التابع.

وتكون العلاقة ما بين المتغيرين علاقة عكسية اذا أدت الزيادة بالمتغير المستقل الى نقصان بالمتغير التابع أو أدى النقصان بالمتغير المستقل الى زيادة في المتغير التابع. ويشير الشكل (١٢ - ١) الى ان العلاقة ما بين كمية الانتاج والتكاليف هي علاقة طردية وكما هو متوقع. ان الجزء (أ) من الشكل (١٢ - ٢) يبين حالة تكون فيها العلاقة ما بين المتغيرين س و ص هي علاقة عكسية. كما انه ليس من الضروري أن تظهر جميع لوحات الانتشار علاقات اما عكسية أو طردية. ان بعض لوحات الانتشار قد تظهر عدم وجود ارتباط كما هو الحال في الجزء (جـ) من الشكل (١٢ - ٢)، والذي يعني أن التغير في المتغير المستقل ليس له تاثير على المتغير التابع.

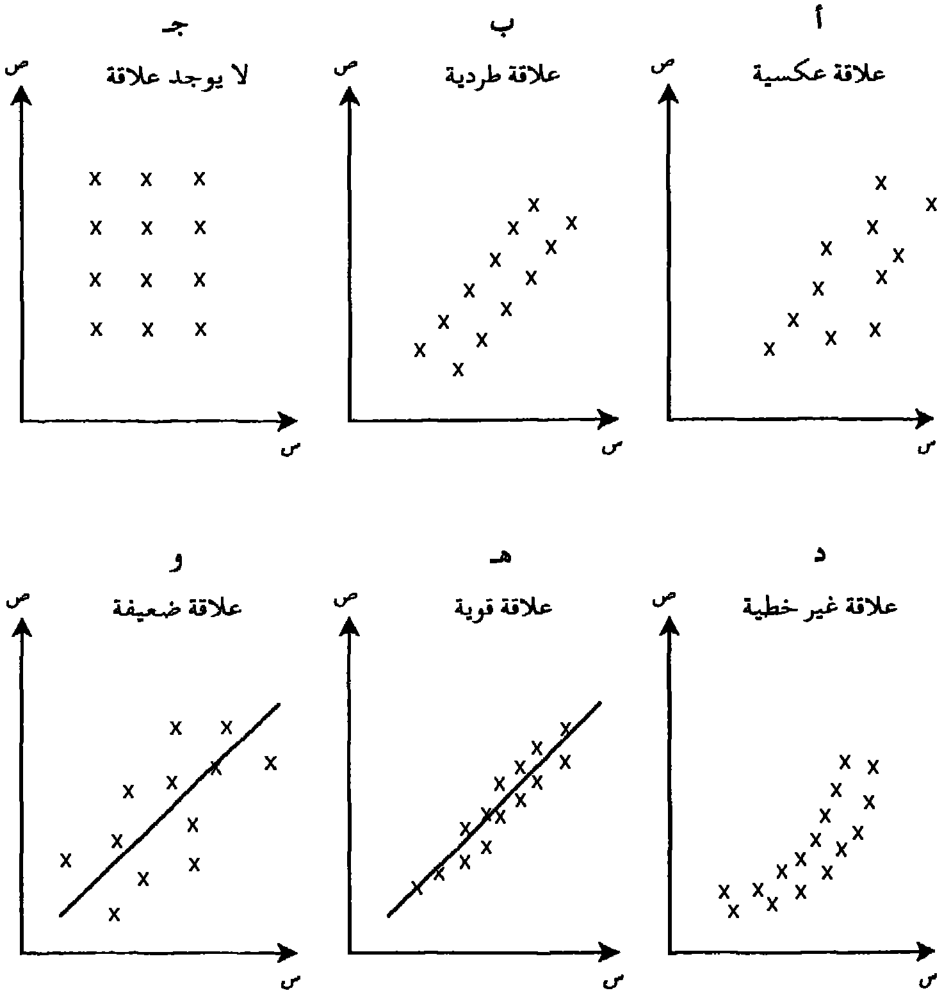
الشكل (١٢ - ١)

لوحة الانتشار لكمية الانتاج والتكاليف



شكل (١٢ - ٢)

لوحات انتشار لعلاقات مختلفة ما بين المتغيرين س و ص



ب ـ هل العلاقة ما بين المتغيرين علاقة خطية أم غير خطية ؟

تكون العلاقة ما بين المتغيرين علاقة خطية اذا كان الخط المستقيم قادر على تزويد تمثيل مناسب لمتوسط العلاقة ما بين المتغيرين. وتكون العلاقة غير خطية ما بين المتغيرين س و ص اذا لم تتمكن النقاط الموجودة على لوحة الانتشار من تشكيل خط مستقيم لتمثيل متوسط العلاقة. والشكل (١٢ - ١) يقترح ان العلاقة ما بين كمية الانتاج والتكاليف هي علاقة خطية. كما ان الجزء (د) من الشكل (١٢ - ٢) يوضح حالة تكون فيها العلاقة ما بين المتغيرين س و ص هي علاقة غير خطية.

جـ ـ هل العلاقة ما بين المتغيرين قوية أم ضعيفة ؟

تكون العلاقة ما بين المتغيرين س و ص نسبياً قوية اذا كانت النقاط تقع على قرب من خط معدل العلاقة (Average Line Relationship). فعلى سبيل المثال يوضح الجزء (هـ) من الشكل (١٢ - ٢) حالة تكون فيها العلاقة ما بين المتغيرين س و ص علاقة قوية. لذلك يستطيع الشخص الباحث في هذه الحالة التنبؤ بقيم المتغير التابع (ص) بناءاً على قيم المتغير المستقل (س) وبدقة لوقوع جميع النقاط بالقرب من الخط. وعلى العكس من ذلك يوضح الجزء (و) من الشكل (١٢ - ٢) حالة تكون فيها العلاقة ما بين المتغيرين س و ص علاقة ضعيفة بحيث انه لا يستطيع الشخص الباحث من تقدير قيم المتغير التابع (ص) بناءاً على قيم المتغير المستقل (ص) بسبب تناثر النقاط عن خط معدّل العلاقة في لوحة الانتشار. ويكون الارتباط تام ما بين المتغيرين في حالة وقوع جميع النقاط في لوحة الانتشار على خط معدل العلاقة.

ـ تحليل الارتباط Correlation Analysis

يهتم تحليل الارتباط بمدى قوّة العلاقة ما بين متغيرين. فكما لاحظنا في الجزء (هـ) والجزء (و) في الشكل (١٢ - ٢)، فان بعض العلاقات ما بين المتغيرات تكون اقوى أو أضعف من بعض العلاقات الأخرى. فعلى سبيل المثال ان حجم القدم اليسرى للشخص له علاقة قوية بحجم القدم اليمنى لنفس الشخص. وعلى الجهة الأخرى، يوجد هناك بعض العلاقة ما بين حجم المؤسسة والقدرة على التكيّف مع تقنية جديدة. لذلك يعتبر تحليل الارتباط مهم ومكمّل لتحليل الانحدار. فتحليل الانحدار يصف نوع العلاقة ما بين متغيرين، بينما تحليل الارتباط يصف قوة هذه العلاقة.

ـ أهداف تحليل الانحدار والارتباط

Aims of Regression and Correlation Analysis

يوجد هناك أربعة أهداف رئيسية بالنسبة لتحليل الانحدار والارتباط. وهذه الأهداف هي :

أ ـ يزود تحليل الانحدار تقديراً للمتغير التابع بناءاً على قيم المتغير المستقل. فاذا أرادت المنشأة تقدير تكاليف انتاج أربعة طن من المخرجات بالشهر من الشكل (١٢ - ١)، فان تحليل الانحدار يزود هذا النوع من التقدير بالاعتماد على خط الانحدار. ان خط الانحدار (والذي غالباً ما يُشتق باستخدام طريقة أقل المربعات والتي سوف نشرحها فيما بعد) يستطيع تقدير متوسط قيمة (Mean Value) المتغير التابع (ص) لكل قيمة من قيم المتغير المستقل (س). لذلك يستطيع خط الانحدار في حالة المنشأة من تقدير متوسط قيمة التكاليف لكل قيمة من قيم كميات الانتاج.

ب ـ يزود تحليل الانحدار قياسات للأخطاء التي يمكن أن تحدث باستخدام خط الانحدار لتقدير المتغير التابع. فعلى سبيل المثال، فانه من المفيد جداً في حالة المنشأة معرفة

مقدار الثقة التي يمكن أن يضعها شخص في تقدير التكاليف بالاعتماد على خط الانحدار. وللإجابة على هذا السؤال، فقد قام الاحصائيون ببناء فترات الثقة والتي سوف نتناولها فيما بعد.

جـ- يستطيع تحليل الانحدار تقدير التأثير الناتج على متوسط قيمة المتغير التابع (ص) بسبب تغير قيمة المتغير المستقل (س) بمقدار وحدة واحدة. فعلى سبيل المثال : ممكن ان يكون هناك اهتمام في حالة المنشأة عند ادارتها بالتكلفة الحدية والتي هي عبارة عن الزيادة في التكلفة الكلية نتيجة زيادة المخرجات بمقدار طن واحد. فاذا كانت العلاقة ما بين كمية الانتاج والتكاليف هي علاقة خطية، فان ميل خط الانحدار يساوي التكلفة الحدية للمنشأة. لذلك يمكن من خلال تحليل الانحدار تقدير ميل خط الانحدار واختبار الفرضية ذات العلاقة بقيمة الميل. وسوف نتناول هذا الموضوع فيما يتبع من هذا الفصل.

د- يزود تحليل الانحدار تقديراً لقوّة العلاقة ما بين المتغيرين س و ص. إن معامل الارتباط ومعامل التحديد هما مقياسان يستخدمان بشكل عام لهذا العرض. وسوف نتناول ذلك فيما يتبع من هذا الفصل.

- نموذج الانحدار الخطي Linear Regression Model

يُعرّف النموذج على انه تمثيل مبسط للعالم الحقيقي. وفي هذا الجزء فاننا نصف النموذج بمجموعة من الافتراضات والتي تعتمد عليها تحليل الانحدار.

يطلق على التوزيع الاحتمالي للمتغير التابع (ص) بناءاً على قيم معروفة للمتغير المستقل (س) بالتوزيع الاحتمالي المشروط للمتغير (ص) والذي يرمز له بالرمز ح (ص ا س). حيث ان :

ص هي قيمة المتغير التابع

س هي قيم محددة للمتغير المستقل.

ويفترض تحليل الانحدار الافتراضات التالية حول التوزيع الاحتمالي المشروط للمتغير ص. وهذه الافتراضات هي :

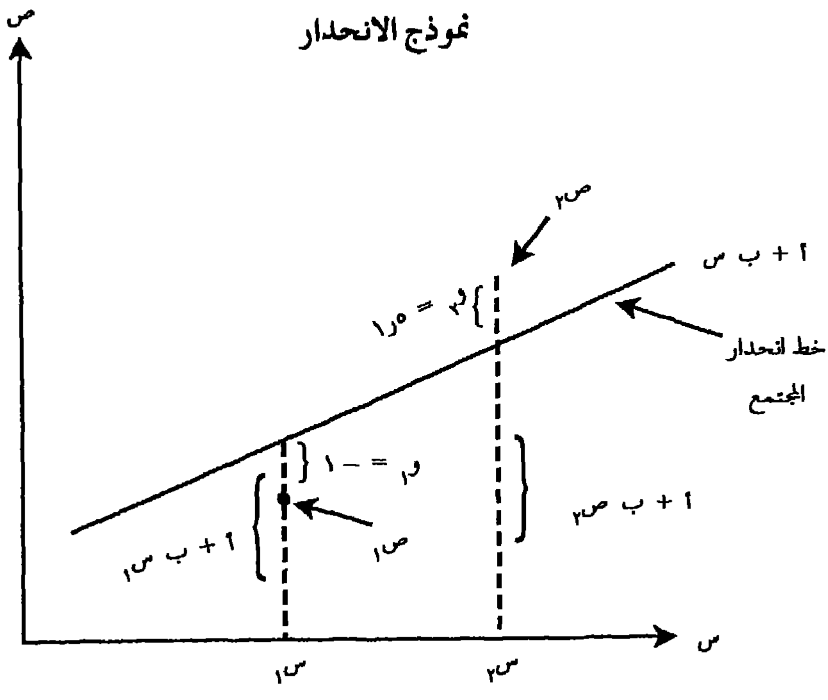
أ- يفترض تحليل الانحدار أن متوسط قيمة المتغير ص، بناءاً على قيم معلومة للمتغير س، هو اقتران خطي للمتغير س. بمعنى آخر، يفترض تحليل الانحدار ان متوسط قيمة المتغير التابع (ص) عبارة عن اقتران خطي للمتغير المستقل (س). وبعبارة أخرى، يفترض تحليل الانحدار وقوع الأوساط للتوزيعات الاحتمالية المشروطة على خط مستقيم. كما ان معادلة هذا الخط هي

$\mu_{(ص ، س)}$ = أ + ب س

والشكل (١٢-٣) يوضح حالة مثل هذا النوع. ويطلق على هذا الخط المستقيم بخط الانحدار للمجتمع أو خط الانحدار الصحيح.

شكل (١٢ - ٣)

نموذج الانحدار



ب - يفترض تحليل الانحدار ان الانحراف المعياري للتوزيع الاحتمالي المشروط هو نفسه وبغض النظر عن قيمة المتغير المستقل. أي ان الانحراف المعياري للتوزيع الاحتمالي للتكلفة عندما تكون قيمة كمية الانتاج هي (٢) طن هو نفس الانحراف المعياري للتوزيع الاحتمالي للتكلفة عندما تكون قيمة كمية الانتاج (٩) طن.

جـ- يفترض تحليل الانحدار ان قيم المتغير التابع (ص) مستقلة عن بعضها البعض. فعلى سبيل المثال : اذا كانت احد المشاهدات تقع تحت وسط التوزيع الاحتمالي المشروط، فان هذا لن يؤثر على احتمال وقوع أحد المشاهدات الأخرى تحت وسط التوزيع الاحتمالي المشروط. وفي مثال المنشأة : اذا كانت تكاليف انتاج أحد الأشهر اقل من المعدّل، فإنه من الممكن ان تكون تكاليف انتاج الشهر اللاحق تحت المعدّل على فرض ثبات عوامل بيئة العمل لشهر آخر.

د - يفترض تحليل الانحدار ان التوزيع الاحتمالي المشروط هو توزيعاً طبيعياً. بالواقع لا تتطلب جميع جوانب تحليل الانحدار هذا الافتراض. كما يجدر بملاحظة هنا ان تحليل الانحدار يعتبر المتغير ص فقط كمتغير عشوائي. كما يفترض تحليل الانحدار ثبات قيم المتغير المستقل. لهذا فان استخدام تحليل الانحدار في تقدير قيمة المتغير التابع بناءاً على المتغير المستقل يجعل القيمة المقدرة للمتغير التابع هي مجال للخطا، اما قيمة المتغير المستقل تكون معروفة. ففي مثال المنشأة، اذا أُستخدم تحليل الانحدار لتقدير تكاليف انتاج (٤) طن من المخرجات، فان القيمة المقدرة للتكاليف هي مجال للخطا فقط، اما بالنسبة لكمية الانتاج فهي معروفة وبدقة.

ان الافتراضات الأربعة معاً لتحليل الانحدار تتضمن ما يلي :

$$ص_ر = أ + ب س_ر + و_ر$$

حيث ان $ص_ر$ هي القيمة الملاحظة ر من المتغير التابع. $س_ر$ هي القيمة الملاحظة

ر من قيم المتغير المستقل . $و_ر$ هو مقدار الخطأ.

إن وجود مقدار الخطأ يؤدي الى وقوع القيم الملاحظة للمتغير ص ر حول خط الانحدار للمجتمع وليس عليه. اذا كانت قيمة $و_١$ = −١ (قيمة مقدار الخطأ)، فان قيمة $ص_١$ سوف يقع تحت خط انحدار المجتمع، والشكل (١٢ - ٣) يوضح ذلك. أما اذا كانت قيمة $و_٢$ = ٥ر١، فهذا يعني أن قيمة $ص_٢$ سوف تقع فوق خط انحدار المجتمع. كما ان تحليل الانحدار يفترض استقلال قيم ($و_ر$) عن بعضها البعض.

- خط انحدار العينة Sample Regression Line

ان تحليل الانحدار يتطلب الحصول على معادلة رياضية لخط الانحدار لوصف معدل العلاقة ما بين المتغير التابع والمتغير المستقل. فاذا تمّ استخدام جميع عناصر مجتمع الدراسة في حساب معادلة خط الانحدار، فان الخط الناتج يُسمّى بمعادلة خط الانحدار للمجتمع. أما اذا لم يتمكن الشخص الباحث من ايجاد معادلة خط الانحدار للمجتمع فانه سوف يقوم بتقدير ذلك الخط بمعادلة خط الانحدار للعينة من خلال استخدام عناصر العينة التي تم اختيارها من ذلك المجتمع. لذلك نلاحظ ان حساب معادلة خط الانحدار للمجتمع تعتمد على جميع عناصر المجتمع، أما معادلة خط الانحدار للعينة فانها تعتمد على جزء أو عينة من مجتمع الدراسة فقط. ان التعبير العام لمعادلة خط الانحدار للمجتمع هو

$$ص = أ + ب س$$

أما التعبير العام لمعادلة خط الانحدار للعينة هو

$$\hat{ص} = أ + ب س$$

حيث ان :

١) $\hat{ص}$ = القيمة المتوقعة للمتغير التابع باستخدام معادلة خط الانحدار.

٢) أ و ب هي قيم تقديرية لـ أ و ب على التوالي.

بما ان معادلة خط الانحدار تعني ضمنياً ان قيمة صˆ = أ اذا كانت قيمة س = صفر، فان هذا يعني ان القيمة أ هي قيمة المتغير ص والتي عندها تقطع معادلة خط الانحدار محور الصادات. لذلك غالباً ما يطلق على قيمة الثابت أ بنقطة تقاطع خط الانحدار مع محور الصادرات. أما قيمة ب والتي تمثل ميل خط الانحدار فانها تعمل على قياس التغير في قيمة ص المتوقعة نتيجة زيادة قيمة المتغير المستقل بمقدار وحدة واحدة.

يمثل الشكل (١٢ ـ ٤) خط الانحدار التقديري للبيانات ذات العلاقة بالتكاليف والانتاج للمنشأة التي تم التطرق لها في بداية هذا الفصل. ان معادلة خط الانحدار هي:

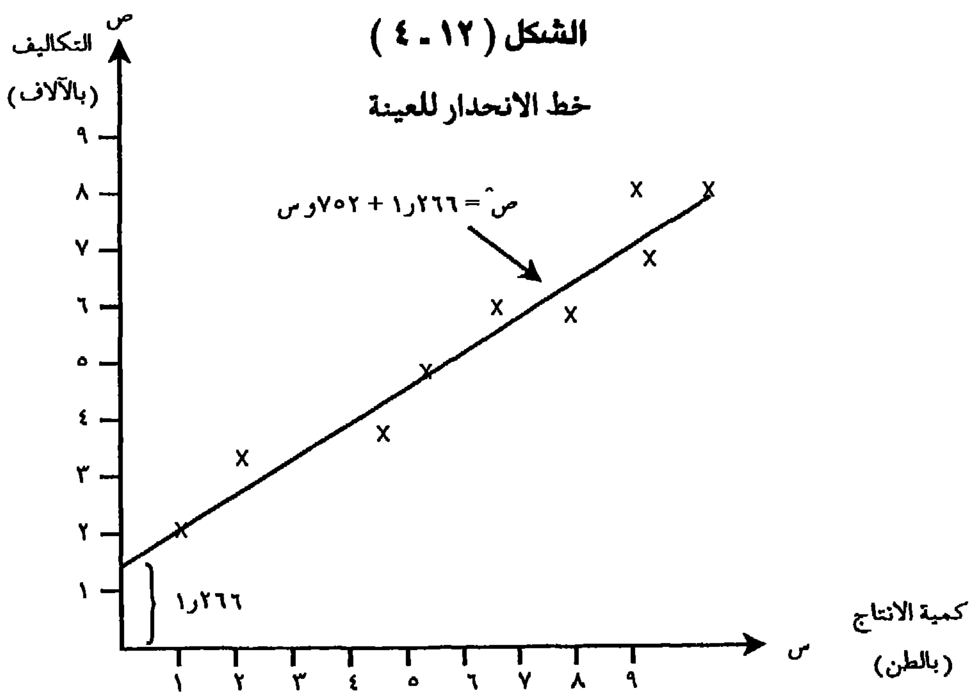
صˆ = ١٫٢٦٦ + ٠٫٧٥٢ س

حيث ان : ص هي التكلفة الشهرية بآلاف الدنانير.

س هي كمية الانتاج الشهرية بالطن.

الشكل (١٢ ـ ٤)

خط الانحدار للعينة



ان القيمة (٢٦٦ر١) في معادلة خط الانحدار للعينة هي قيمة أ والتي هي قيمة تقديرية لـ أ. أما بالنسبة للقيمة (٧٥٢ر) فهي قيمة ب والتي هي قيمة تقديرية لـ **ب**.

ان الاهتمام هنا ليس بكيفية أو آلية حساب معادلة خط الانحدار للعينة والذي سوف يتم شرحه بالتفصيل بالجزء التالي، ولكن الاهتمام هنا سوف يكون بكيفية تفسير معادلة خط الانحدار.

قبل البدء بأي شيء فانه لا بد من ملاحظة الفرق ما بين **ص** و صˆ. ان القيمة **ص** تشير الى قيمة حقيقية للتكلفة الشهرية، بينما قيمة صˆ تشير الى القيمة المحسوبة أو القيمة التقديرية للتكلفة الشهرية بناءاً على معادلة خط الانحدار للعينة. فعلى سبيل المثال، ان الصف الأول في الجدول (١٢ - ١) يبين ان التكلفة الشهرية الحقيقية لانتاج طن واحد هي (٢) ألف دينار. لذلك ان قيمة **ص** = ٢ ألف دينار عندما تكون قيمة س = ١ طن. بالمقارنة ان معادلة خط الانحدار تشير الى ان صˆ = ٢٦٦ر١ + ٧٥٢، (١) = ٠١٨ر٢ ألف دينار عندما تكون قيمة س = (١) طن. بمعنى آخر، ان القيمة التقديرية المتوقعة للتكاليف الشهرية تساوي (٠١٨ر٢) ألف دينار عندما تكون كمية الانتاج طن واحد، بينما القيمة الحقيقية للتكاليف تحت نفس الظروف (تكلفة الشهر الأول) هي (٢) ألف دينار.

انه من المهم جداً أن تتوفر المقدرة على ايجاد وتفسير كل من نقطة تقاطع خط الانحدار ومحور الصادرات وكذلك ميل الخط. ما هي نقطة تقاطع خط الانحدار مع محور الصادرات في حالة المنشأة وماذا تعني؟ إن نقطة التقاطع هي (٢٦٦ر١) وهذا يعني انه اذا كان الانتاج الشهري للمنشأة صفر طن فان التكلفة الشهرية سوف تكون (٢٦٦ر١) دينار كما يظهر في الشكل (١٢ - ٤)، وتمثل هذه القيمة (٢٦٦ر١) قيمة المتغير التابع والتي عندها يقطع خط الانحدار محور الصادرات. ما هو ميل خط الانحدار في هذه الحالة؟ ان ميل خط الانحدار هو (٧٥٢ر) ألف دينار. ان هذا يعني ان الزيادة التقديرية بالتكاليف الشهرية اذا زاد الانتاج بمقدار طن واحد هي ٧٥٢ دينار.

ـ طريقة أقل المربعات Least Squares Method

يهتم هذا الجزء بوصف كيفية حساب خط معدل الانحدار. ولتوضيح ذلك، افرض ان شخص ما يريد تقدير خط الانحدار لتمثيل العلاقة ما بين التكلفة والانتاج في الشكل (١٢ ـ ١). بما ان معادلة خط الانحدار هي :

ص̂ = أ + ب س ،

فان عملية التقدير للمعادلة تتم عن طريق ايجاد القيم الرقمية بالنسبة للثوابت أ و ب. ان عدد البدائل بالنسبة للقيم التي يمكن ان تأخذها الثوابت أ و ب هي بدائل غير منتهية (Infinite Number). ان احد البدائل للقيم التي يمكن ان يأخذها كل من أ و ب كما يظهر في الجزء (أ) من الشكل (١٢ ـ ٥) هو ان قيمة أ = ($\frac{١}{٢}$) و ب = (١). أما البديل الثاني الذي يمكن ان تأخذه قيم كل من أ و ب كما يظهر في الجزء (ب) من الشكل (١٢ ـ ٥) هو ان قيمة أ = (٠) وقيمة ب = (١). كما ان احد البدائل الأخرى لقيم أ و ب كما يظهر في الجزء (جـ) من الشكل (١٢ ـ ٥) هو ان قيمة أ = (–١) وقيمة ب = (١). ان السؤال الذي يطرح نفسه الآن هو السؤال التالي :

ما هو البديل الأمثل من بين هذه البدائل في وصف العلاقة ؟

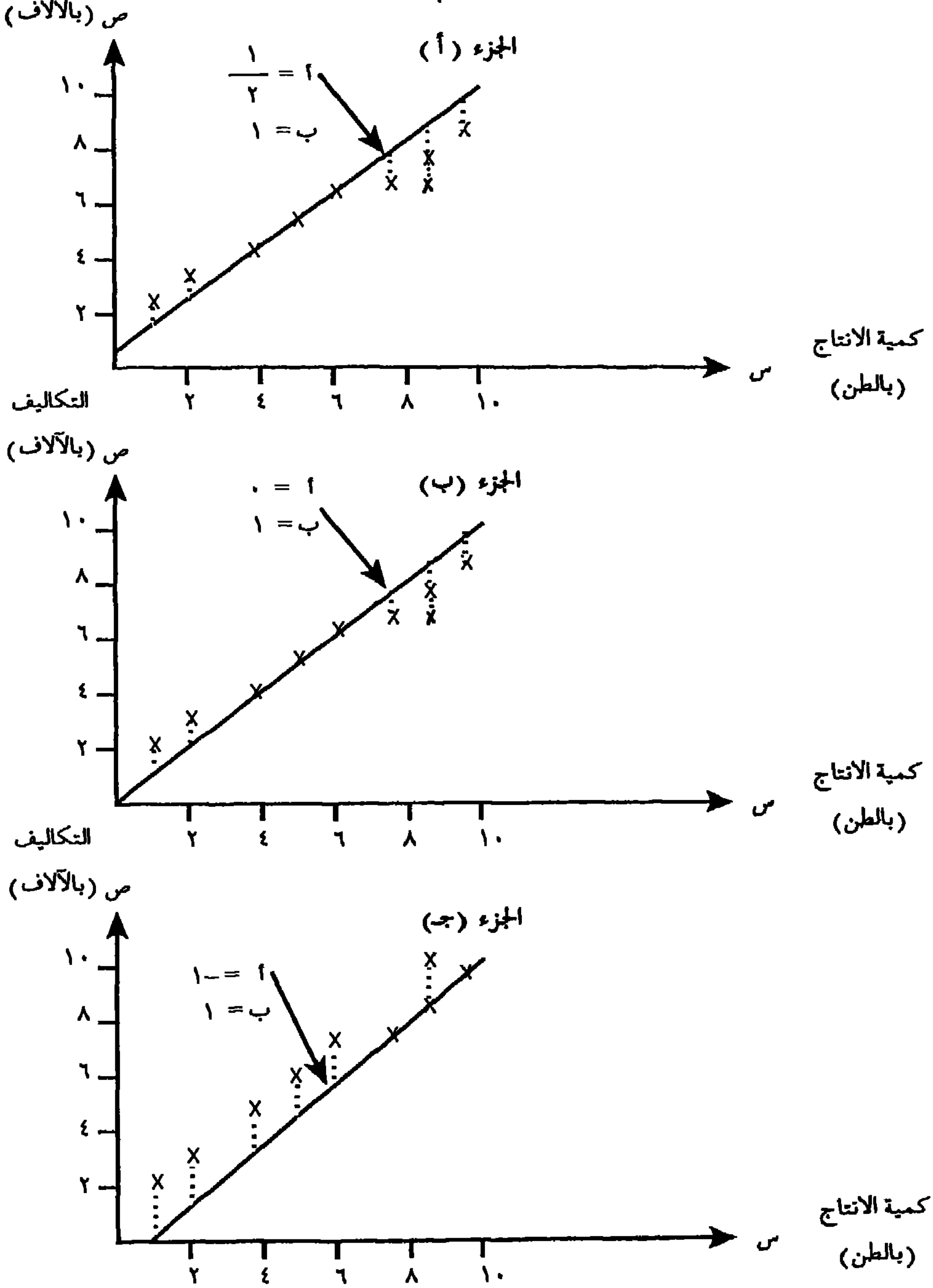
ان طريقة أقل المربعات تعطي الاجابة على هذا السؤال من خلال الطريقة التالية : يجب اعتبار كل بديل من بدائل معادلة خط الانحدار (كل قيمة متوقعة من قيم أ او ب) من خلال قياس انحراف كل مشاهدة من مشاهدات العينة عن الخط. لهذا، ان قياس الانحراف في الجزء (أ) من الشكل (١٢ ـ ٥) قد تمّ من خلال الخطوط العمودية المتقطعة ما بين الخط والنقاط. بعد ذلك وبناءاً على طريقة أقل المربعات نقوم بتربيع هذه الفروق ومن ثم جمعها معاً. إن مجموع مربع الانحرافات للمشاهدات عن خط الانحدار للجزء أ من الشكل (١٢ ـ ٥) هو [(٥ر)^{٢} + (٥ر)^{٢} + (–٥ر)^{٢} + (–٥ر)^{٢} + (–٥ر)^{٢} + (–٣/٢)^{٢} + (–٣/٢)^{٢} + (–٥ر)^{٢} + (–٣/٢)^{٢}] = (٨ر٢٥). أما بالنسبة للجزء (ب) من الشكل (١٢ ـ ٥) فان مجموع مربع انحرافات المشاهدات عن

خط الانحـدار هو $[\ (١)^{٢} + (١)^{٢} + (٠)^{٢} + (٠)^{٢} + (٠)^{٢} + (-١)^{٢} + (-١)^{٢} + (٠)^{٢} + (-١)^{٢}\ ] = (٥)$. أما بالنسبة للجزء (جـ) من الشكل (١٢ ـ ٥) فان مجموع مربع الانحرافات للمشاهدات عن خط الانحدار هو $[\ (٢)^{٢} + (٢)^{٢} + (١)^{٢} + (١)^{٢} + (١)^{٢} + (٠)^{٢} + (١)^{٢} + (٠)^{٢} + (٠)^{٢}\ ] = (١٢)$. اذن وبناءاً على طريقـة أقل المربعات فاننا سوف نختار الخط الذي مجموع مربع انحرافات المشاهدات عنه أقل ما يمكن.

انه من المنطق ان نختار الخط الذي يقلل مجموع مربع انحرافات بيانات العينة عنه. لماذا، لانه كلما زاد مجموع مربع انحرافات المشاهدات عن خط معدل العلاقة كلما قلت مقدرة الخط في تقدير أو تمثيل البيانات. لذلك يجب على أي شخص ان يقلل من مجموع مربع انحرافات المشاهدات عن خط معدل العلاقة اذا أراد ان يحصل على خط قادر على تمثيل البيانات بأكبر قدر ممكن. وبالنظر الى الشكل (١٢ ـ ٥) نجد ان الخط في الجزء (جـ) لا يمثل البيانات كما في الجزء (أ)، كما ان الخط في الجزء (أ) لا يمثل البيانات كما في الجزء (ب). ان هذه الحقيقة موضحة من خلال الفروق في مربع انحرافات المشاهدات عن الخطوط. وكما هو متوقع، ان مجموع مربع انحرافات المشاهدات في الجزء (جـ) أكبر من مجموع مربع الانحرافات في الجزء (أ)، كما ان مجموع مربع انحرافات المشاهدات في الجزء (أ) أكبر من مجموع مربع الانحرافات في الجزء (ب) وكما هو واضح في الشكل (١٢ ـ ٥) .

الشكل (١٢ - ٥)

بدائل لقيم أ و ب



ورياضيـاً، اذا كـان $\text{س}_\text{ر}$ و $\text{ص}_\text{ر}$ يمثل الزوج المرتب ر من المشـاهدات ذات العـلاقـة بالمتغيرات المستقلة والمتغير التابع، فانه يمكن ايجاد قيم كل من أ و ب والتي تقلل من مجموع مربع انحرافات المشاهدات عن خط الانحدار باستخدام المعادلات التالية :

$$\sum_{\text{ر}=1}^{\text{ن}} \text{ص}_\text{ر} = \text{ن} \times \text{أ} + \text{ب} \sum_{\text{ر}=1}^{\text{ن}} \text{س}_\text{ر}$$

$$\sum_{\text{ر}=1}^{\text{ن}} \text{س}_\text{ر}\text{ص}_\text{ر} = \text{أ} \sum_{\text{ر}=1}^{\text{ن}} \text{س}_\text{ر} + \text{ب} \sum_{\text{ر}=1}^{\text{ن}} \text{س}_\text{ر}^2$$

حيث ان ن هو عدد قيم المتغير أما ($\text{س}_\text{ر}$) و ($\text{ص}_\text{ر}$) والذي بناءاً عليهن تم حساب خط الانحدار للعينة. وبحل المعادلتين اعلاه بالنسبة لـ أ و ب (والذي غالباً ما يستخدما لتقدير أ و ب بطريقة اقل المربعات) فاننا سوف نحصل على المعادلتين الهامتين التاليتين :

$$\text{ب} = \frac{\sum_{\text{ر}=1}^{\text{ن}} (\text{س}_\text{ر} - \bar{\text{س}})(\text{ص}_\text{ر} - \bar{\text{ص}})}{\sum_{\text{ر}=1}^{\text{ن}} (\text{س}_\text{ر} - \bar{\text{س}})^2} \quad .... \quad (\text{١٢-أ})$$

$$\text{أ} = \bar{\text{ص}} - \text{ب}\,\bar{\text{س}} \quad .... \quad (\text{١٢-ب})$$

وتمثل القيمة ب في المعادلة (١٢ -أ) معامل الانحدار التقديري ومن وجهة نظر حسابية، فانه غالباً ما يفضل استخدام معادلة أخرى لـ(ب) غير المعادلة (١٢ -أ)، علماً بأن كلا المعادلتين يؤديان الى نفس النتيجة. ان هذه المعادلة هي :

$$\text{ب} = \frac{\sum_{\text{ر}=1}^{\text{ن}} \text{س}_\text{ر}\text{ص}_\text{ر} - (\sum_{\text{ر}=1}^{\text{ن}} \text{س}_\text{ر})(\sum_{\text{ر}=1}^{\text{ن}} \text{ص}_\text{ر})}{\sum_{\text{ر}=1}^{\text{ن}} \text{س}_\text{ر}^2 - (\sum_{\text{ر}=1}^{\text{ن}} \text{س}_\text{ر})^2}$$

ولايجاد معادلة الانحدار لوصف العلاقة ما بين كمية الانتاج والتكاليف بالنسبة للجدول (١٢ - ١)، فاننا سوف نقوم بتشكيل الجدول (١٢ - ٢) لحساب قيم كل من Σ سر صر ، Σ س^{٢}ر ، Σ سر ، Σ صر . وبالاعتماد على الحسابات في الجدول (١٢ - ٢) ، نجد ما يلي :

$$ب = \frac{٩(٣١٩) - (٥٠)(٤٩)}{٩(٣٤٠) - (٥٠)^{٢}}$$

$$ب = \frac{٢٨٧١ - ٢٤٥٠}{٣٦٠ - ٢٥٠٠}$$

$$ب = \frac{٤٢١}{٥٦٠}$$

ب = ٧٥٢ر

لذلك تكون قيمة ب باستخدام طريقة اقل المربعات هي (٧٥٢ر) من الألف دينار والتي هي نفس النتيجة التي ظهرت بالجزء السابق. بمعنى آخر، اذا زاد الانتاج بمقدار طن واحد، فان ذلك سوف يتبعه زيادة تقديرية بالتكاليف قدرها حوالي (٧٥٢) دينار.

جدول (١٢ - ٢)

$س_ر$	$ص_ر$	$س_ر^٢$	$س_ر ص_ر$	$ص_ر^٢$
١	٢	١	٢	٤
٢	٣	٤	٦	٩
٤	٤	١٦	١٦	١٦
٨	٧	٦٤	٥٦	٤٩
٦	٦	٣٦	٣٦	٣٦
٥	٥	٢٥	٢٥	٢٥
٨	٨	٦٤	٦٤	٦٤
٩	٨	٨١	٧٢	٦٤
٧	٦	٤٩	٤٢	٣٦
٥٠	٤٩	٣٤٠	٣١٩	٣٠٣

بعد حساب قيمة ب، فانه لا بد من حساب قيمة أ لايجاد معادلة خط الانحدار والتي تؤدي الى تقليل مجموع مربع الانحرافات. وباستخدام المعادلة (١٢ - ب) .

أ = صَ − ب سَ حيث ان :

صَ = الوسط الحسابي لقيم المتغير ص .

سَ = الوسط الحسابي لقيم المتغير المستقل س .

لايجاد قيمة أ ، فانه لا بد أولاً من ايجاد قيمة كل من سَ و صَ .

$$\text{إن قيمة س َ} = \frac{٥٠}{٩}$$

$$\text{س َ} = ٥\text{ر}٥٥٦$$

$$\text{أما قيمة ص َ} = \frac{٤٩}{٩}$$

$$\text{ص َ} = ٥\text{ر}٤٤٤$$

وبعد التعويض بالمعادلة (١٢ ـ ب)، نجد ان قيمة أ هي :

$$\text{أ} = ٥\text{ر}٤٤٤ - ٧٥٢\text{ر}(٥\text{ر}٥٥٦)$$

$$\text{أ} = ٥\text{ر}٤٤٤ - ٤\text{ر}١٧٨$$

$$\text{أ} = ١\text{ر}٢٦٦$$

وهذا يعني ان قيمة أ هي (١ر٢٦٦) دينار. تذكر هنا ان هذه القيمة هي نفس القيمة المعطاة للثابت أ في الجزء السابق.

وبعد الحصول على قيم كل من الثابت أ والثابت ب، فانه من السهل جداً تحديد معدل العلاقة للعينة ما بين الانتاج والتكاليف بالنسبة للمنشأة. وهذه العلاقة هي :

$$\hat{\text{ص}} = ١\text{ر}٢٦٦ + ٧٥٢\text{ر س}$$

حيث ان قيمة ص تم قياسها بآلاف الدنانير، بينما القيمة س تم قياسها بالطن. وكما نعلم، فانه يطلق على هذا الخط بخط الانحدار للعينة أو خط انحدار المتغير ص بناءاً على س. كما ان هذا الخط هو نفس الخط الذي تم تمثيله بيانياً في الشكل (١٢ ـ ٤).

ان خط الانحدار من هذا النوع قد يكون له أهمية كبيرة في الحياة العملية بالنسبة للاداريين والاقتصاديين على حد سواء. فعلى سبيل المثال : افرض ان ادارة

المنشأة تريد تقدير تكاليف الانتاج الشهرية اذا قرّرت انتاج (٤) طن في كل شهر. فباستخدام معادلة خط الانحدار لهذه المنشأة، فانه يمكن تقدير التكاليف كما يلي :

ص = ١ر٢٦٦ + ٧٥٢ر (٤)

= ٤ر٢٧٤

بما ان التكاليف تم قياسها على أساس ألف دينار، فان هذا يعني ان التكاليف الكلية المقدرة سوف تكون ٤ر٢٧٤ دينار . ولتوضيح الفكرة ، دعنا نقوم بحل مثال آخر.

مثال (١) : يريد احصائي تقدير العلاقة ما بين الدخل السنوي للعائلة وحجم الادخار. فاذا تم الحصول على البيانات من تسع عائلات.

الدخل السنوي (ألف) : ١٢، ١٣، ١٤، ١٥، ١٦، ١٧، ١٨، ١٩، ٢٠

الادخار السنوي (ألف) : ٠، ١ر ، ٢ر ، ٢ر ، ٥ر ، ٥ر ، ٦ور ، ٧ر ، ٨ر

احسب معادلة خط الانحدار البسيط بطريقة أقل المربعات اذا علمت ان الادخار هو المتغير التابع والدخل السنوي هو المتغير المستقل.

الحل : افرض ان $س_ر$: الدخل السنوي بآلاف الدنانير.

$ص_ر$: الادخار السنوي بآلاف الدنانير.

سوف نقوم الآن بترتيب البيانات في جدول (١٢ - ٣) لحساب قيم كل من

$\sum س_ر$ ، $\sum س^{٢}_ر$ ، $\sum ص_ر$ ، $\sum س_ر ص_ر$.

جدول (١٢ - ٣)

$س_ر$	$ص_ر$	$\sum س_ر^٢$	$\sum س_ر ص_ر$	$\sum ص_ر^٢$
١٢	٠	١٤٤	٠	٠
١٣	١ر	١٦٩	١ر٣	٠١ر
١٤	٢ر	١٩٦	٢ر٨	٠٤ر
١٥	٢ر	٢٢٥	٣ر٠	٠٤ر
١٦	٥ر	٢٥٦	٨ر٠	٢٥ر
١٧	٥ر	٢٨٩	٨ر٥	٢٥ر
١٨	٦ر	٣٢٤	١٠ر٨	٣٦ر
١٩	٧ر	٣٦١	١٣ر٣	٤٩ر
٢٠	٨ر	٤٠٠	١٦ر٠	٦٤ر
١٤٤	٣ر٦	٢٣٦٤	٦٣ر٧	٢ر٠٨

لايجاد معادلة خط الانحدار، فانه لا بد من ايجاد قيمة كل من ب وقيمة ا، لايجاد ب فاننا سوف نستخدم العلاقة التالية :

$$ب = \frac{ن \sum س_ر ص_ر - (\sum س_ر)(\sum ص_ر)}{ن \sum س_ر^٢ - (\sum س_ر)^٢}$$

وبالاعتماد على الحسابات في الجدول (١٢ - ٢) نجد ما يلي :

ب = [٩(٦٣ر٧) − (١٤٤) (٣ر٦)] / [٩(٢٣٦٤) − (١٤٤)^٢]

ب = [٥٧٣ر٣ − ٥١٨ر٤] / [٢١٢٧٦ − ٢٠٧٣٦]

ب = ٥٤ر٩ / ٥٤٠

ب = ر١٠١٧

ان هذا يعني ان قيمة ب باستخدام طريقة أقل المربعات هي (١٠١ر٧) دينار. بمعنى آخر، كلما زاد دخل الأسرة أو العائلة بمقدار الف دينار، كلما زاد الادخار بمقدار (١٠١ر٧) دينار.

ولايجاد قيمة ا، فاننا سوف نستخدم المعادلة التالية :

ا = صَ − ب سَ

ان قيمة صَ = Σ صر / ن

صَ = ٣ر٦ / ٩

صَ = ر٤

اما قيمة سَ = Σ سر / ن

سَ = ١٤٤/٩

سَ = ١٦

اذن قيمة ا = ر٤ − (ر١٠١٧) (١٦)

أ = ٤ر – ١ر٦٢٧٢

أ = ١ر٢٢٧٢

وهذا يعني ان قيمة أ هي (–٢ر١٢٢٧) دينار. وهذا يعني انه اذا كان دخل الأسرة السنوي (صفر) دينار، فان الأسرة بحاجة الى اقتراض المبلغ (٢ر١٢٢٧) لقضاء الحاجات الأساسية.

وعليه فان معادلة خط الانحدار هي :

صَ = – ١ر٢٢٧٢ + ١٠١٧ر س

علماً بانه تم قياس كل من الدخل والادخار بآلاف الدنانير.

ـ خصائص التقدير بطريقة أقل المربعات

ان الهدف النهائي من قيام الشخص الباحث بالعمليات الحسابية هو الحصول على معادلة خط الانحدار للمجتمع. فعلى سبيل المثال ان هدف اي منشأة هو معرفة خط الانحدار للمجتمع والذي يربط تكاليف المنشأة بكمية الانتاج ـ خط الانحدار الذي يعتمد على جميع مشاهدات التكاليف والانتاج. ولكن من الصعب جداً في بعض الأحيان القيام بحساب هذا الخط لعدم وجود جميع عناصر المجتمع. لذلك فان أفضل شيء بالنسبة للشخص الباحث هو ايجاد خط الأنحدار بالنسبة لعينة من مجتمع الدراسة ومن ثم استخدامه لتقدير خط الانحدار للمجتمع. لقد تم تعريف القيم الاحصائية أ و ب كقيم تقديرية لمعالم خط انحدار المجتمع **أ** و **ب** . وسواء كان التوزيع الاحتمالي المشروط للمتغير طبيعي أم غير طبيعي، فان هذه التقديرات تتمتع بالخصائص التالية :

١ ـ عدم التحيز **Unbiasedness**

انه يمكن ملاحظة ان القيم المحسوبة لـ أ و ب بالنسبة للعينة باستخدام طريقة اقل المربعات هي قيم تقديرية غير منحازة للقيم أ و ب بالنسبة للمجتمع . بمعنى آخر ، اذا كان لا بد من سحب عينة تلوا الأخرى ، وتم حساب قيمة كل من أ و ب لكل عينة باستخدام طريقة اقل المربعات، فان الوسط الحسابي لقيم أ سوف يساوي القيمة الحقيقية لـ أ ، كما ان الوسط الحسابي لقيم ب سوف يساوي القيمة الحقيقية لـ ب .

٢ ـ الكفاءة **Efficiency**

الانحراف المعياري بالنسبة للتقديرات غير المنحازة أ و ب باستخدام طريقة اقل المربعات غالباً ما يكون اقل ما يمكن. كما أنه من أسباب تفضيل استخدام طريقة اقل المربعات في تقدير قيم كل من أ و ب هو تقليل مجموع مربع انحرافات المشاهدات عن خط معدل العلاقة للعينة. بمعنى آخر، ان القيم التقديرية أ و ب هما اكثر التقديرات كفاءة لمثل هذا النوع. ولقد اثبتت هذة النتيجة عن طريق نظرية قوس و ماركوف (**Gauss - Markov Theorem**) .

٣ ـ الثبات على مبدأ أو الانسجام **Consistency**

لقد تم ايجاد ان الثابت أ بالنسبة للعينة هو تقدير ثابت أو منسجم مع الثابت أ بالنسبة للمجتمع وكذلك نفس الشيء بالنسبة للثابت ب . بمعنى آخر ، كلما زاد حجم العينة واصبح اكبر واكبر، فان قيمة أ بالنسبة للعينة سوف تنطبق على قيمة أ بالنسبة للمجتمع، كما ان قيمة ب بالنسبة للعينة سوف تنطبق على قيمة ب بالنسبة للمجتمع.

و بناءاً على هذة الخصائص المميزة جداً . فان القيم التقديرية لـ أ و ب

باستخدام طريقة اقل المربعات هي القيم التقديرية التي استخدمها الاحصائيون لتقدير الثوابت بالنسبة لمجتمع الدراسة أ و ب .

الخطأ المعياري للتقدير Standarad Error of Estimates

لقد بيّنا في الجزء السابق كيف ان تحليل الانحدار يزود تقادير مختلفة للمتغير التابع بناءاً على قيم معطاة للمتغير المستقل (وكان هذا أول أهداف تحليل الانحراف والارتباط) . أما الآن فاننا سوف نقوم بوصف كيفية تزويد تحليل الانحراف لمقاييس الأخطاء التي يمكن أن ترافق اجراءات التقدير (الهدف الثاني من أهداف تحليل الانحدار والارتباط) . انه لمن الضروري أن نتذكر أولاً أننا قد قمنا بافتراض مساواة الانحراف المعياري للتوزيع الاحتمالي المشروط للمتغير التابع وبغض النظر عن قيمة المتغير المستقل. ان الهدف من استخدام مقياس الانحراف المعياري والذي يرمز له بالرمز (σe) هو لقياس كمية الانتشار عن خط انحدار المجتمع. فاذا كانت قيمة (σe) كبيرة، فان هذا يعني ان الانتشار كبير. أما اذا كانت قيمة (σe) صغيرة، فهذا يعني ان الانتشار قليل.

ان المقياس الاحصائي المستخدم لتقدير σe هو الانحراف المعياري للتقدير والذي يرمز له بالرمز (Se) والمعرّف كما يلي :

$$Se = \sqrt{\frac{\sum_{ر=1}^{ن} (ص_ر - \hat{ص}_ر)^2}{ن - 1}} ،$$

حيث أن ١) $ص_ر$ هي القيمة ر من قيم المتغير التابع.

٢) ن هي حجم العينة.

٣) $\hat{ص}_ر$ هي القيمة التقديرية لـ $ص_ر$ من خط الانحدار الذي معادلته هي

$\hat{ص}_ر = أ + ب س_ر$.

انه من الواضح جداً ان تزداد قيمة (Se) كلما زادت كمية الانتشار عن خط الانحدار للعينة. فاذا لم يكن هناك انتشار عن خط الانحدار للعينة على الاطلاق (أي ان جميع النقاط واقعة على خط الانحدار) فان قيمة (Se) سوف تكون صفراً لان قيمة $ص_ر$ سوف تكون نفس قيمة $\hat{ص}_ر$. ولكن اذا كان هناك انتشار كبير عن خط الانحدار، فان هذا يعني ان قيمة $ص_ر$ سوف تختلف اختلافاً كبيراً عن قيمة $\hat{ص}_ر$ وبالتالي سوف تكون قيمة (Se) كبيرة.

ان المعادلة التي تستخدم بشكل مستمر في حساب الخطأ المعياري للتقدير بسبب السهولة في التطبيق والمكافئة للمعادلة الأولى هي :

$$Se = \sqrt{\frac{\sum_{ر=1}^{ن} ص_ر^2 - أ \sum_{ر=1}^{ن} ص_ر - ب \sum_{ر=1}^{ن} س_ر ص_ر}{ن - 1}}$$

ولتوضيح استخدام المعادلة اعلاه، دعنا نعود الى نفس مثال المنشأة التي تريد وصف العلاقة ما بين كمية الانتاج والتكاليف. بالرجوع الى الجدول (١٢ - ٢) نجد ان

$\sum ص_ر^2 = ٣٠٣$ ، $\sum ص_ر = ٤٩$ ، $\sum س_ر ص_ر = ٣١٩$.

كما أن قيمة أ = ١ر٢٦٦ ، وقيمة ب = ر٧٥٢

وبالتعويض بالمعادلة نجد ان قيمة (Se) هي

$$Se = \sqrt{\frac{(٣٠٣) - (١ر٢٦٦)(٤٩) - (ر٧٥٢)(٣١٩)}{٩ - ١}}$$

$$= \sqrt{\frac{٣٠٣ - (٦٢ر٠٣٤) - (٢٣٩ر٨٨٨)}{٧}}$$

$$= \sqrt{\frac{\text{١ر٠٧٨}}{\text{٧}}}$$

$$= \sqrt{\text{ر١٥٤}}$$

$$= \text{ر٣٩٢}$$

لذلك ، اذا كنا نعلم القيم الحقيقية لكل من **أ** و **ب** بالنسبة لمجتمع الدراسة ، فان الانحراف المعياري للخطأ بالتقدير بالاعتماد على خط الانحدار الصحيح للمجتمع سوف يكون حوالي (٣٩٢ر) بالألف من الدينار أو ٣٩٢ دينار.

مثال (٢) : بالاعتماد على البيانات الواردة في المثال رقم (١)، اذا أراد الاقتصادي حساب الخطأ المعياري في التقدير، فما هي تلك القيمة ؟

الحل : بالرجوع الى البيانات الواردة في الجدول (١٣ ـ ٢) نجد أن :

$\sum \text{ص}_\text{ر}^{\text{٢}} = \text{٢ر٠٨}$ ، $\sum \text{ص}_\text{ر} = \text{٣ر٦}$ ، $\sum \text{س}_\text{ر}\,\text{ص}_\text{ر} = \text{٦٣ر٧}$.

كما ان قيمة أ = ١ر٢٢٧٢ ، قيمة ب = ر١٠١٧

بالتعويض بالمعادلة التالية، نجد ان قيمة الخطأ المعياري بالتقدير هي :

$$S_e = \sqrt{\frac{\sum_{\text{ر}=\text{١}}^{\text{ن}} \text{ص}_\text{ر}^{\text{٢}} - \text{أ} \sum_{\text{ر}=\text{١}}^{\text{ن}} \text{ص}_\text{ر} - \text{ب} \sum_{\text{ر}=\text{١}}^{\text{ن}} \text{س}_\text{ر}\,\text{ص}_\text{ر}}{\text{ن - ١}}}$$

$$S_e = \sqrt{\frac{\text{٢ر٠٨} - (\text{١ر٢٢٧٢})(\text{٣ر٦}) - (\text{ر١٠١٧})(\text{٦٣ر٧})}{\text{٩ - ٢}}}$$

$$= \sqrt{\frac{٢٫٠٨ - (٤٫٤١٧٩) - (٦٫٤٧٨٣)}{٧}}$$

$$= \sqrt{٠٫٠٠٢٨}$$

$$= ٠٫٠٥٣$$

لذلك، ان الانحراف المعياري للخطأ بالتقدير بالاعتماد على القييم الحقيقية لخط الانحدار الصحيح هو حوالي (٠٫٠٥٣) بالألف من الدينار أو (٥٣) دينار.

ـ تقدير الوسط المشروط Estimating the Conditional Mean

ان الهدف من هذا الجزء هو بيان كيفية تقدير الوسط المشروط للمتغير ص. ولغاية هذا الجزء، فاننا نفترض ان كل من خط الانحدار والتقدير يعتمد على طريقة أقل المربعات في تقدير قيم كل من **أ** و **ب** الحقيقية. ولغايات التحديد ، دعنا نعود الى مثال المنشأة التي ترغب في وصف العلاقة ما بين كمية الانتاج والتكاليف. افرض ان مدراء المنشأة يرغبون في تقدير متوسط التكاليف الشهرية للمنشأة في حالة تحديد كمية الانتاج الشهرية لأن تكون (٤) طن.

بما أن خط الانحدار الحقيقي للمجتمع مجهول، فان أفضل شيء قد تقوم به المنشأة هو تقدير خط الانحدار الصحيح بخط الانحدار للعينة. بمعنى آخر، لتقدير تكاليف انتاج (٤) طن، فاننا سوف نستخدم معادلة خط الانحدار للعينة وكما يلي :

$$\hat{ص} = أ + ب\,س$$

$$\hat{ص} = ١٫٢٦٦ + ٠٫٧٥٢\,س_ر$$

$$\hat{ص}(٤) = ١٫٢٦٦ + ٠٫٧٥٢\,(٤)$$

$$= ١٫٢٦٦ + ٣٫٠٠٨$$

$$= ٤٫٢٧٤$$

وبشكل عام، لتقدير الوسط المشروط للمتغير عندما تكون قيمة المتغير المستقل تساوي س* ، فاننا نستخدم العلاقة التالية :

$$أ + ب س^*$$

أما بالنسبة لفترة الثقة للوسط المشروط للمتغير ص اذا كانت قيمة المتغير المستقل هي س* هي :

$$(أ + ب س^*) \pm (t_{\frac{\alpha}{2}})(Se)\sqrt{\frac{1}{ن} + \frac{(س^* - \bar{س})^2}{\sum_{ر=1}^{ن} س_ر^2 - ن\bar{س}^2}}$$

فاذا كان التوزيع الاحتمالي المشروط للمتغير المستقل توزيعاً طبيعياً، فان احتمال أن تتضمن هذه الفترة الوسط المشروط للمتغير ص هو (١ - α). لاحظ ان هذه الفترة سوف تتسع أكثر وأكثر كلما بعدت قيمة س* عن سَ. ان اتساع الفترة نتيجة لبعد س* عن سَ هو منطقي لأن الخطأ الناتج في تقدير ميل خط الانحدار سوف يؤدي الى وقوع خطأ أكبر في تقدير موقع خط الانحدار.

ولتوضيح استخدام العلاقة بالنسبة لفترة الثقة، دعنا نقوم بتقدير متوسط التكاليف المتحققة على المنشأة في حالة الرغبة في تحديد حجم الانتاج الشهري بـ(٤) طن وبشكل دوري. ان فترة الثقة الـ ٩٥٪ لمتوسط التكاليف هي :

$$٤٫٢٧٤ \pm (٢٫٣٦٥)(٫٣٩٢)\sqrt{\frac{1}{٩} + \frac{(٤ - ٥٫٥٥٦)^2}{٣٤٠ - ٩(٥٫٥٥٦)^2}}$$

$$= ٤٫٢٧٤ \pm (٫٩٢٧١)\sqrt{\frac{1}{٩} + \frac{(-١٫٥٥٦)^2}{٣٤٠ - ٢٧٧٫٨٢}}$$

$$= \text{٤ر٢٧٤} \pm (\text{ر٩٢٧١}) \sqrt{\frac{\text{١}}{\text{٩}} + \frac{\text{٢ر٤٢١١}}{\text{٦٢ر١٨}}}$$

$$= \text{٤ر٢٧٤} \pm (\text{ر٩٢٧١})(\text{ر٣٨٧٢})$$

$$= \text{٤ر٢٧٤} \pm \text{ر٣٥٩}$$

لاحظ ان قيمة (t) اعتمدت على ($\propto$/٢) ودرجات الحرية (٩ – ٢ = ٧). لذلك، اذا أرادت المنشأة تقدير النقطة الواقعة على خط انحدار المجتمع والمقابلة لكمية الانتاج الـ (٤) طن الشهرية، فان فترة الثقة الـ ٩٥٪ لهذا الوسط المشروط هي من (٩١٥ر٣ الى ٦٣٣ر٤) دينار.

ان الهدف من المثال التالي هو لزيادة توضيح كيفية حساب فترة الثقة للوسط المشروط.

مثال (٣) : أوجد فترة الثقة الـ ٩٥٪ لمتوسط حجم الادخار لعائلة دخلها السنوي (٢٠ر٠٠٠) دينار من بين مجموعة من العائلات بالاعتماد على البيانات الواردة في المثال رقم (١).

الحل : بالاعتماد على البيانات الواردة في الجدول (١٢ ـ ٢) نلاحظ ما يلي :

$$\text{ا} = -\text{١ر٢٢٧٢} \text{ ، ب} = \text{ر١٠١٧} \text{ ، } t_{(\text{ر٠٢٥ ، د.ح}=\text{٧})} = \text{٢ر٣٦٥}$$

$$S_e = \text{ر٠٥٣} \text{ ، ن} = \text{٩} \text{ ، } \bar{\text{س}} = \text{١٦} \text{ ، } \text{س}_{\text{ر}}^{\text{٢}} = \text{٢٣٦٤} \text{ ، س}^{*} = \text{٢٠ الف.}$$

بالتعويض بالمعادلة فإن فترة الثقة الـ ٩٥٪

$$= -\text{١ر٢٢٧٢} + (\text{ر١٠١٧})(\text{٢٠}) \pm (\text{٢ر٣٦٥})(\text{ر٠٥٣}) \sqrt{\frac{\text{١}}{\text{٩}} + \frac{(\text{٢٠} - \text{١٦})^{\text{٢}}}{\text{٢٣٦٤} - \text{٩}(\text{١٦})^{\text{٢}}}}$$

$$= -٢٢٧٢ر١ + ٠٣٤ر٢ \pm (٢٥٣ر١) \sqrt{\frac{١}{٩} + \frac{(٤)^٢}{٢٣٦٤ - ٢٣٠٤}}$$

$$= ٠٦٨ر٨ \pm (٢٥٣ر١) \sqrt{\frac{١}{٩} + \frac{١٦}{١٠}}$$

$$= ٠٦٨ر٨ \pm (٢٥٣ر١)(٦١٥ر)$$

$$= ٠٦٨ر٨ \pm ٧٧٠ر$$

وهذا يعني ان فترة الثقة الـ ٩٥٪ لمتوسط حجم الادخار المشروط هي من ٢٩٨ر٧ الف دينار الى ٨٣٨ر٨ الف دينار، أو ٧٢٩٨ دينار الى ٨٨٣٨ دينار.

التنبؤ بالقيم الفردية للمتغير ص Predicting an Individual Value of Y

غالباً ما يستخدم خط انحدار العينة لتقدير قيم المتغير ص. فاذا أراد مدراء المنشأة توقّع تكاليف الشهر القادم في حالة انتاج (٤) طن، فانه بالامكان استخدام معادلة خط الانحدار البسيط للعينة للقيام بذلك التوقع. وعليه تكون القيمة المتوقعة لتكاليف الشهر القادم هي (٤٢٧٤) دينار. بشكل عام، يمكن لأي شخص الاعتماد على معادلة خط الانحدار للعينة في عملية توقّع قيمة فردية للمتغير ص = أ + ب س*

أما بالنسبة لفترة الثقة لقيمة المتغير ص المتوقعة اذا كانت قيمة المتغير هي س* فهي :

$$(أ + ب س^*) \pm (t \frac{\alpha}{٢}) (Se) \sqrt{\frac{ن + ١}{ن} + \frac{(س^* - \bar{س})^٢}{\sum_{ر=١}^{ن} س_ر^٢ - ن \bar{س}^٢}}$$

فاذا كان التوزيع الاحتمالي المشروط للمتغير الطبيعي هو توزيعاً طبيعياً، فان احتمال أن تقع القيمة الصحيحة للمتغير ص ضمن فترة الثقة هي ٩٥٪. ان مقارنة فترة الثقة للقيم الفردية للمتغير ص مع فترة الثقة للوسط المشروط، نلاحظ ان هذه الفترة أوسع من فترة الثقة الأخرى. ان هذا بالطبع منطقي لأن الخطأ العيني في تقدير القيمة الفردية للمتغير التابع سوف يكون أكبر من الخطأ العيني لقيمة الوسط المشروط بالنسبة للمتغير التابع.

ولتوضيح استخدام فترة الثقة، افرض ان ادارة المنشاة تهتم بتوقع التكلفة التي سوف تتحقق للشهر القادم اذا كانت كمية الانتاج المطلوبة هي (٤) طن. بالاعتماد على العلاقة، تكون فترة الثقة الـ ٩٥٪ للتكلفة :

$$= (١ر٢٦٦ + ٧٥٢ر (٤)) \pm (٢ر٣٦٥) (٣٩٢ر) \sqrt{\frac{١٠}{٩} + \frac{(٤ - ٥ر٥٠٦)^{٢}}{٦٢ر٢}}$$

$$= ٤ر٢٧٤ \pm (٩٢٧٠ر) \sqrt{\frac{١٠}{٩} + \frac{(- ١ر٥٠٦)^{٢}}{٦٢ر٢}}$$

$$= ٤ر٢٧٤ \pm (٩٢٧٠ر) (١ر٠٧٢٣)$$

$$= ٤ر٢٧٤ \pm ٩٩٤ر$$

ولقد تمّ حساب قيمة ($t \frac{\partial}{٢}$) بناءاً على درجات حرية قدرها (٩ – ٢ = ٧). لهذا، اذا كانت المنشاة مهتمة في تقدير التكلفة لشهر معين اذا كانت كمية الانتاج هي (٤) طن، فان فترة الثقة الـ ٩٥٪ لهذه التكلفة هي من ٣ر٢٨٠ دينار الى ٥ر٢٦٨ دينار.

وقبل ان ننتهي من موضوع فترة الثقة لكل من الوسط المشروط والقيم الفردية للمتغير ص، فانه لا بد من التذكير بالافتراضات ذات العلاقة مرة أخرى. أولاً : لقد تمّ

افتراض ان انحدار المجتمع هو خطي بمعنى آخر، ان الوسط المشروط للمتغير التابع هو اقتران خطي للمتغير المستقل. ثانياً : لقد تم افتراض ان التوزيع الاحتمالي المشروط للمتغير التابع هو توزيعاً طبيعياً والانحراف المعياري هو نفسه بغض النظر عن قيمة المتغير المستقل. فاذا تمّ نقض اي من هذين الافتراضين، فان هذا سوف يؤدي الى وقوع درجة معينة من الخطأ في النتائج.

مثال (٤) : بالاعتماد على البيانات المتوفرة في المثال رقم (١)، ما هو الادخار المتوقع لعائلة تم اختيارها عشوائياً من بين العائلات ذات الدخل (٢٠ر٠٠٠) دينار، ثم أوجد فترة الثقة الـ ٩٥٪ لهذا التوقع او التنبؤ.

الحل : نلاحظ من البيانات الواردة في الجدول (١٢ - ٢) ما يلي :

أ = – ١ر٢٢٧٢ ، ب = ١٠١٧ر ، t (٠٢٥ر ، ٧) = ٢ر٣٦٥

Se = ٠٥٣ر ، ن = ٩ ، سَ = ١٦ ، $\sum$ س$_{ر}^{٢}$ = ٢٣٦٤

س* = ٢٠ الف.

أ) التنبؤ = – ١ر٢٢٧٢ + ١٠١٧ر (٢٠)

= – ١ر٢٢٧٢ + ٢ر٠٣٤

= ٨٠٦٨ر

ب) فترة الثقة الـ ٩٥٪ هي

فترة الثقة = –١ر٢٢٧٢ + ١٠١٧ر (٢٠) ± (٢ر٣٦٥) (٠٥٣ر) $\sqrt{\frac{١٠}{٩} + \frac{(٢٠ - ١٦)^{٢}}{٢٣٦٤ - ٩(١٦)^{٢}}}$

= ٨٠٦٨ر ± (٢ر٣٦٥) (٠٥٣ر) (١ر١٧٤)

= ٨٠٦٨ر ± ١٤٧٢ر

أو = ٨٠٦٨ر إلى ٩٥٤ر

لهذا، فان فترة الثقة الـ ٩٥٪ بالنسبة للادخار لعائلة تمّ اختيارها عشوائياً من بين العائلات من ذوي الدخل (٢٠٫٠٠٠) دينار هي من ٦٥٩٦ر الى ٩٥٤ر ديناراً.

ـ معامل التحديد Coefficient of Determination

لقد بيّنا في الفصول السابقة كيفية حساب معادلة خط الانحدار. فبعد ايجاد خط الانحدار، فان اهتمام الباحث يتوجه نحو معرفة ما اذا كان هذا الخط ممثلاً للبيانات أم لا وكما يظهر في الشكل (١٢ - ٦). فمن الشكل يظهر ان خط الانحدار في الجزء (ب) من الشكل يزود تمثيل افضل للبيانات مقارنة من خط الانحدار في الجزء (أ) من نفس الشكل. فكيف يمكن قياس درجة تمثيل خط الانحدار للبيانات ؟

ان أول خطوة تجاه الحل لهذا السؤال تتطلب شرح معنى التغيّر، والذي يشير الى مجموع الانحرافات. ان مجموع الانحرافات للمتغير التابع ص يساوي

$$\sum_{ر=١}^{ن} (ص_ر - \bar{ص})^٢$$

بمعنى آخر، ان مجموع الانحرافات يساوي مجموع مربع الانحرافات لقيم ص عن وسطها.

ولقياس درجة تمثيل خط الانحدار لمجموعة من البيانات، فاننا سوف نقوم بتقسيم التغير الكلي للمتغير التابع الى قسمين : (١) التغيّر الذي يمكن تفسيره بواسطة خط الانحدار. (٢) التغيّر الذي لا يمكن تفسيره بواسطة خط الانحدار. ولتقسيم التغير الكلي لهذه الطريقة، فانه لا بد من ملاحظة الآتي بالنسبة للمشاهدة (ر)

$$(ص_ر - \bar{ص}) = (ص_ر - \hat{ص}_ر) + (\hat{ص}_ر - \bar{ص})$$

علماً بأن $\hat{ص}_ر$ هي قيمة $ص_ر$ المتوقعة بالاعتماد على خط الانحدار. بمعنى آخر وكما يظهر في الشكل (١٢ - ٧)، فانه يمكن تقسيم المسافة ما بين ص ر وقيمة الوسط

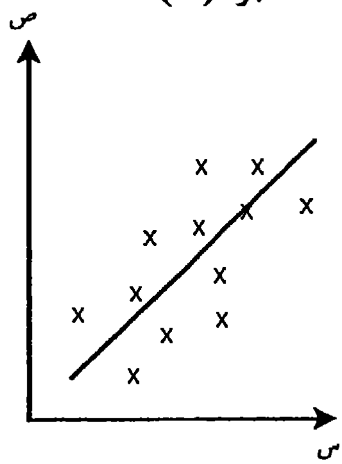
للمتغير ص الى قسمين : المسافة ما بين صر والنقطة الواقعة على خط الانحدار والتي تقع مباشرة تحت (فوق) صر، والمسافة الواقعة مباشرة تحت النقطة الواقعة على خط الانحدار والوسط صَ.

فاذا قمنا الآن بتربيع كلا الطرفين بالنسبة لمعادلة التغير الكلي وجمعنا النتيجة بالنسبة لجميع قيم (ر)، فاننا سوف نحصل على :

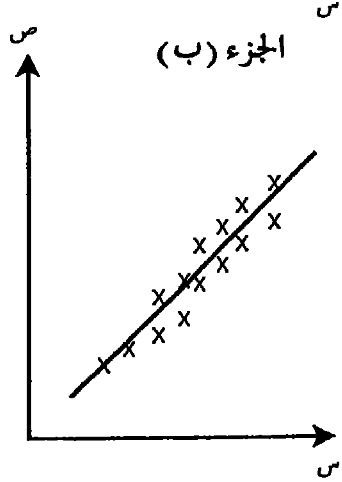
الشكل (١٢ - ٦)

درجة التمثيل لخطي انحدار

الجزء (أ)

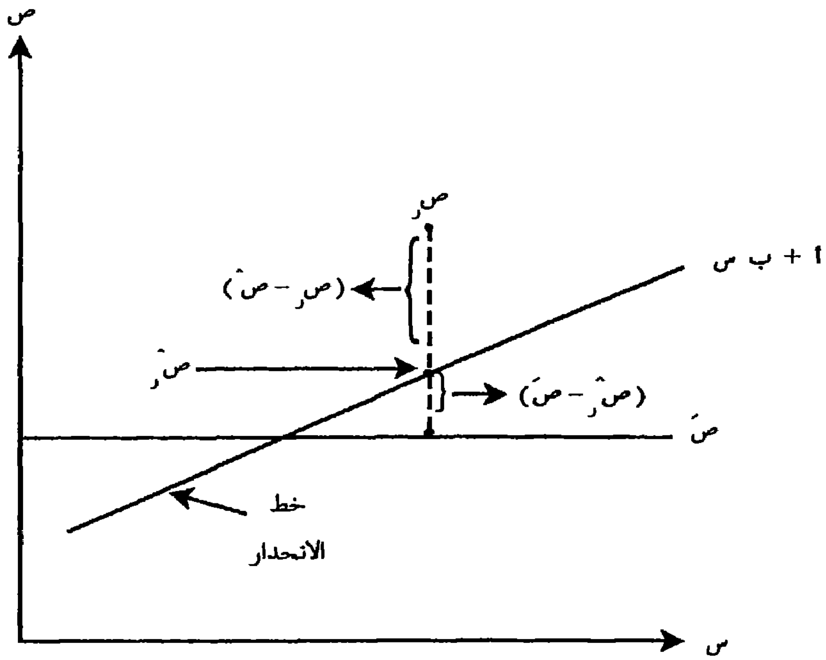


الجزء (ب)



الشكل (١٢ ـ ٧)

تقسيم (ص ر − صَ) الى قسمين : (ص ر − صˆ ر) و (صˆ ر − صَ)



$$\sum_{ر=١}^{ن} (ص_ر - \bar{ص})^٢ = \sum_{ر=١}^{ن} (ص_ر - \hat{ص}_ر)^٢ + \sum_{ر=١}^{ن} (\hat{ص}_ر - \bar{ص})^٢$$

ان الطرف الأيمن للمعادلة يبين التغير الكلي للمتغير التابع. أما بالنسبة للمقدار الأول من الجهة اليسرى للمعادلة فانه يقيس التغير في المتغير التابع الذي لا يمكن تفسيره عن طريق الانحدار. ويعتبر هذا التفسير مناسباً لهذا المقدار لانه يمثل مجموع مربع الانحرافات للبيانات الحقيقية عن خط الانحدار. لذلك، كلما زادت قيمة هذا المقدار كلما قلت درجة تمثيل خط الانحدار للبيانات.

اما المقدار الثاني من الجهة اليسرى للمعادلة فانه يقيس التغير في المتغير التابع الذي يمكن تفسيره عن طريق خط الانحدار. ويعتبر هذا التفسير مناسباً لهذا المقدار والذي يبين مقدار التغير في المتغير التابع بناءاً على خط الانحدار فقط. بعبارة اخرى، ان هذا المقدار يبين كمية التناقص في التغيّر الذي لا يمكن توضيحه بسبب استخدام الانحدار بدلاً من استخدام (صَ) كقيمة تقديرية للمتغير (ص).

ولقياس درجة تمثيل خط الانحدار للبيانات، فقد استخدم الاحصائيون ما يسمى بمعامل التحديد والذي يعرّف كما يلي:

$$ر^{2} = 1 - \frac{\sum_{ر=1}^{ن} (ص_ر - \hat{ص}_ر)^{2}}{\sum_{ر=1}^{ن} (ص_ر - \bar{ص})^{2}}$$

بعبارة أخرى

$$ر^{2} = 1 - \frac{\text{التغير الذي لم يفسّر باستخدام الانحدار}}{\text{مجموع التغير}}$$

$$ر^{2} = \frac{\text{التغير المفسّر باستخدام الانحدار}}{\text{مجموع التغير}}$$

لذلك، يعتبر معامل التحديد مقياساً مناسباً لقياس درجة تمثيل خط الانحدار للبيانات لانه يساوي النسبة من التغير الكلّي بالنسبة للمتغير التابع والتي لم تفسّر عن طريق خط الانحدار.

وبالحياة العملية، ان المعادلة الأكثر ملاءمة لقياس معامل التحديد هي :

$$\text{ر}^2 = \frac{\left[\text{ن} \sum_{\text{ر}=1}^{\text{ن}} \text{س}_\text{ر}\text{ص}_\text{ر} - \left(\sum_{\text{ر}=1}^{\text{ن}} \text{س}_\text{ر}\right)\left(\sum_{\text{ر}=1}^{\text{ن}} \text{ص}_\text{ر}\right)\right]^2}{\left[\text{ن} \sum_{\text{ر}=1}^{\text{ن}} \text{س}_\text{ر}^2 - \left(\sum \text{س}_\text{ر}\right)^2\right]\left[\text{ن} \sum \text{ص}_\text{ر}^2 - \left(\sum \text{ص}_\text{ر}\right)^2\right]}$$

ولتوضيح حسابات معامل التحديد، دعنا نعود لمثال المنشأة التي تريد وصف العلاقة ما بين كمية الانتاج والتكاليف الشهرية. بالاعتماد على البيانات الواردة في الجدول (١٢ ـ ٢) نجد ان جميع المعلومات المطلوبة لحساب معامل التحديد متوفرة في هذا الجدول، حيث ان :

$\text{ن} = ٩$ ، $\sum \text{س}_\text{ر}\text{ص}_\text{ر} = ٣١٩$ ، $\sum \text{س}_\text{ر} = ٥٠$ ، $\sum \text{ص}_\text{ر} = ٤٩$ ،

$\sum \text{س}_\text{ر}^2 = ٣٤٠$ ، $\sum \text{ص}_\text{ر}^2 = ٣٠٣$.

وبالتعويض في معادلة معامل التحديد نجد ان قيمة

$$\text{ر}^2 = \frac{\left[٩(٣١٩) - (٥٠)(٤٩)\right]^2}{\left[٩(٣٤٠) - (٥٠)^2\right]\left[٩(٣٠٣) - (٤٩)^2\right]}$$

$$\text{ر}^2 = \frac{\left[٢٨٧١ - ٢٤٥٠\right]^2}{\left[٣٠٦٠ - ٢٥٠٠\right]\left[٢٧٢٧ - ٢٤٠١\right]}$$

$$\text{ر}^2 = \frac{(٤٢١)^2}{\left[٥٦٠\right]\left[٣٢٦\right]}$$

$$\text{ر}^2 = \frac{١٧٧٢٤١}{١٨٢٥٦٠}$$

$$\text{ر}^2 = ٫٩٧$$

لذلك، فان معامل التحديد ما بين التكلفة والانتاج بالمنشأة هو (٩٧ر). بمعنى آخر، ان خط الانحدار في الشكل (١٢ - ٤) يستطيع تفسير ما نسبته (٩٧٪) من التغير في التكاليف.

كما انه يوجد هناك معادلة أخرى لحساب معامل التحديد والتي تعتبر أحياناً أكثر ملائمة وهي :

$$ر^{2} = \frac{أ \sum_{ر=1}^{ن} ص_{ر} + ب \sum_{ر=1}^{ن} س_{ر} ص_{ر} - \frac{1}{ن}\left(\sum_{ر=1}^{ن} ص_{ر}\right)^{2}}{\sum_{ر=1}^{ن} ص_{ر}^{2} - \frac{1}{ن}\left(\sum_{ر=1}^{ن} ص_{ر}\right)^{2}}$$

ان الايجابية بالنسبة لاستخدام هذه العلاقة مقارنة مع العلاقة الأولى هو امكانية استخدامها في حالة توفر قيم كل من أ و ب. فاذا تمّ استخدام هذه المعادلة لحساب معامل التحديد بالنسبة للمنشأة لقياس درجة تمثيل خط الانحدار للبيانات، فان النتيجة سوف تكون :

$$ر^{2} = \frac{(١ر٢٢٦)(٤٩) + (ر٧٥٢) - (\frac{١}{٩})(٤٩)^{2}}{٣٠٣ - \frac{١}{٩}(٤٩)^{2}}$$

$$ر^{2} = \frac{٦٢ر٠٣٤ + ٢٣٩ر٨٨٨ - ٢٦٦ر٧٧٨}{٣٦ر٢٢٢}$$

$$ر^{2} = \frac{٣٥ر١٤٤}{٣٦ر٢٢٢}$$

$$ر^{2} = ٩٧\ \%$$

اننا نلاحظ ان الاجابة التي حصلنا عليها باستخدام هذه العلاقة هي نفس الاجابة التي حصلنا عليها في الفقرة السابقة.

ـ معامل الارتباط الخطي Linear Correlation Coefficient

ان الهدف أو الغرض من تحليل الارتباط كما تبين في بداية هذا الفصل هو لقياس قوة العلاقة ما بين المتغير (المتغيرات) المستقلة والمتغير التابع. ان الافتراضات التي يحتاجها نموذج تحليل الانحدار هي كما يلي:

١ ـ ان التوزيع العشوائي لكلا المتغيرين المستقل والتابع هو توزيع طبيعي. بينما يفترض تحليل الانحدار ان التوزيع العشوائي للمتغير التابع هو توزيع طبيعي فقط.

٢ ـ ان الانحراف المعياري للمتغير التابع ($\text{ص}_{\text{ر}}$) هو ثابت بالنسبة لجميع قيم المتغير المستقل (س)، كما ان الانحراف للمتغير المستقل ثابت بالنسبة لجميع قيم المتغير التابع.

ان معامل الارتباط غالباً ما يستخدم لقياس قوة العلاقة ما بين متغيرين. ان معامل الارتباط (ر) ببساطة هو الجذر التربيعي لمعامل التحديد ($\text{ر}^{٢}$). وعليه

$$\text{ر} = \sqrt{\text{ر}^{٢}}$$

كما يجب أن تتساوى إشارة معامل الارتباط مع اشارة خط الانحدار. فاذا كان ميل خط الانحدار موجباً (ب > ٠)، فاننا سوف نعتبر قيمة الجذر الموجبة كقيمة لمعامل الارتباط. أما اذا كان ميل خط الانحدار سالباً (ب < ٠)، فاننا سوف نعتبر قيمة الجذر سالبة كقيمة معامل الارتباط. ولتسهيل عملية حساب معامل الارتباط، فاننا سوف نستخدم المعادلة التالية:

$$\text{ر} = \frac{\text{ن} \sum_{\text{ر}=١}^{\text{ن}} \text{س}_{\text{ر}}\text{ص}_{\text{ر}} - \left(\sum_{\text{ر}=١}^{\text{ن}} \text{س}_{\text{ر}}\right)\left(\sum_{\text{ر}=١}^{\text{ن}} \text{ص}_{\text{ر}}\right)}{\sqrt{\text{ن} \sum_{\text{ر}=١}^{\text{ن}} \text{س}_{\text{ر}}^{٢} - \left(\sum \text{س}_{\text{ر}}\right)^{٢}}\ \sqrt{\text{ن} \sum \text{ص}_{\text{ر}}^{٢} - \left(\sum \text{ص}_{\text{ر}}\right)^{٢}}}$$

ان قيم معامل الارتباط تتراوح ما بين ± ١ . بمعنى آخر، لا يمكن ان تكون قيمة معامل الارتباط أكبر من (١) ولا اقل من (−١) . فاذا كانت قيمة ر = (+١)، فهذا يعني وجود علاقة خطية تامة ما بين المتغير المستقل والمتغير التابع ونوع العلاقة طردية. بمعنى آخر، يشبه الحالة في الجزء (أ) من الشكل (١٢ - ٧) . وعلى الجهة الأخرى، اذا كانت قيمة ر = (−١)، فهذا يعني وجود علاقة خطية تامة ما بين المتغير المستقل والتابع ونوع العلاقة عكسية. بمعنى آخر، يشبه الحالة في الجزء (ب) من الشكل (١٢ - ٧) . وفي كل حالة يتكون فيها العلاقة ما بين المتغيرين تامة، فهذا يعني ان خط الانحدار يفسّر جميع التغير في المتغير التابع بسبب وقوع جميع النقاط على خط الانحدار. لماذا تكون قيمة (ر) تساوي اما (+١) أو (−١) اذا كان خط الانحدار يفسر جميع التغير في المتغير التابع؟ لانه تحت هذه الظروف يجب ان تساوي القيمة الفعلية (Actual Value) للمتغير التابع (ص) مع القيمة المحسوبة (Computed Value) عن طريق خط الانحدار (أي ان $\sum(ص_ر - \hat{ص}_ر) = ٠$) بمعنى آخر ، لا يوجد أي تغير في المتغير التابع بحيث لم يُفسر عن طريق خط الانحدار. وهذا يعني ان قيمة معامل التحديد تساوي (١).

أما اذا كانت قيمة (ر = ٠)، فهذا يعني عدم وجود ارتباط خطي على الاطلاق ما بين المتغير المستقل والمتغير التابع. وفي هذه الحالة، سوف تكون القيمة التقديرية لميل خط الانحدار باستخدام طريقة أقل المربعات مساوية للصفر (ب = ٠)، وهذا يعني ان التغير في المتغير المستقل سوف لا يؤثر على المتغير التابع. وتحت هذه الحالة تكون قيمة (ر = ٠) لأن خط الانحدار لا يفسر أي شيء من التغير في المتغير التابع. وبعبارة أخرى، ان $\sum(\hat{ص}_ر - \bar{ص})^2 = ٠$ لأن الوسط الحسابي للمتغير ص يساوي دائماً القيمة المحسوبة عن طريق خط الانحدار. كما ان $\sum(ص_ر - \hat{ص}_ر)^2 = \sum(ص_ر - \bar{ص})^2$ لأن قيمة $\hat{ص}_ر$ تساوي دائماً قيمة $\bar{ص}$. وهذا يعني ان قيمة معامل التحديد ($ر^2$) تساوي (٠)، والذي بدوره يعني ان قيمة معامل الارتباط (ر) تساوي (٠) . وعليه فان افضل تقدير للمتغير التابع في هذه الحالة هو الوسط الحسابي لقيم المتغير التابع ($\bar{ص}$)، لأن

قيمة المتغير المستقل لا تزود أي معلومة مفيدة على هذه القيمة (صَ). والحالة في الجزء (جـ) من الشكل (١٢ - ٧) يشبه هذه الحالة.

ولتوضيح كيفية حساب وتفسير معامل الارتباط العيني، دعنا نقوم بحل المثال التالي :

مثال (٥) : حصل عالم نفس صناعي على علامات الذكاء (IQ) والانتاجية لعشرة عمال، وكانت النتائج كما يلي :

IQ	الانتاجية
١١٠	٥ر٢
١٢٠	٦ر٠
١٣٠	٦ر٣
١٢٦	٥ر٧
١٢٢	٤ر٨
١٢١	٤ر٢
١٠٣	٣ر٠
٩٨	٢ر٩
٨٠	٢ر٧
٩٧	٣ر٢

المطلوب : ١) أوجد معامل الارتباط ما بين علاقات الذكاء والانتاجية.

٢) هل العلاقة طردية أم عكسية ؟

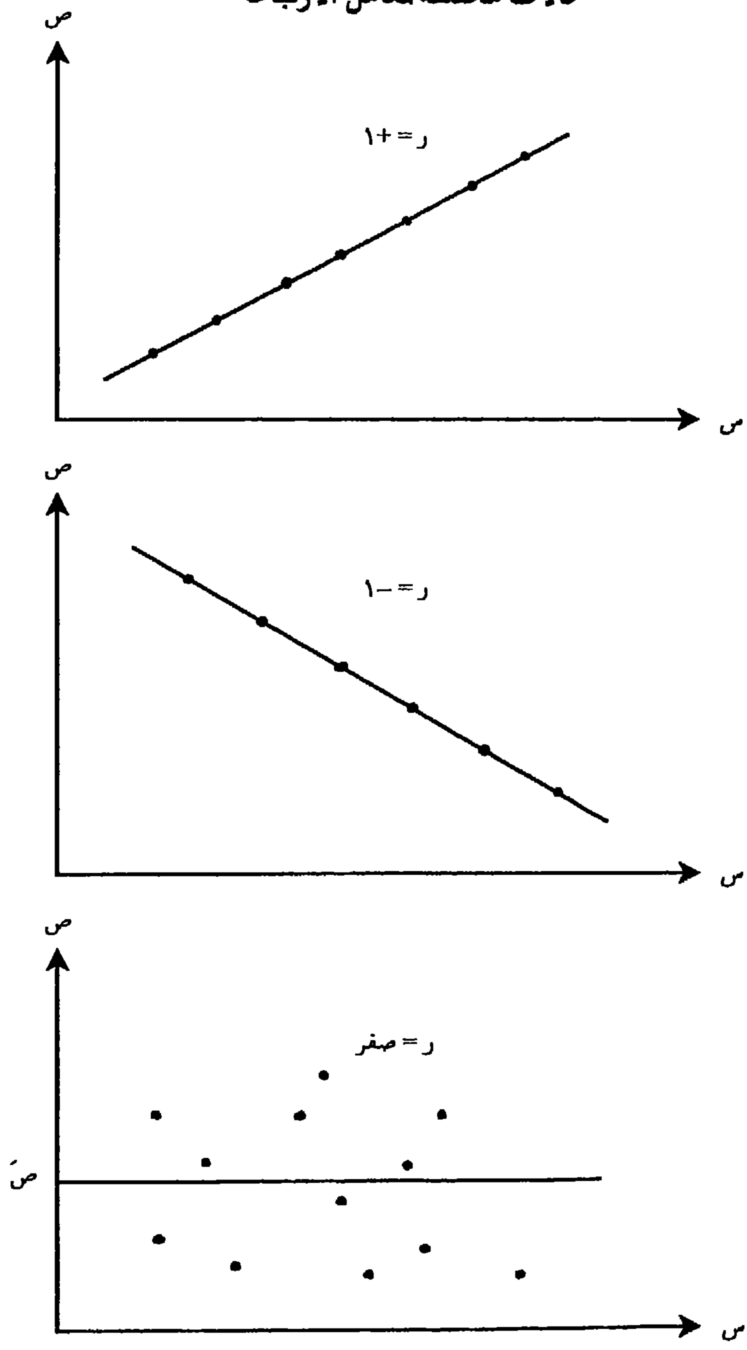
الحل : قبل البدء بحساب معامل الارتباط الخطي، فانه لا بد من تحديد المتغير المستقل والتابع. لذلك نفرض ان :

$س_ر$ = علامة الذكاء للعامل (ر).

$ص_ر$ = انتاجية العامل (ر).

شكل (١٢ ـ ٧)

حالات مختلفة لمعامل الارتباط



دعنا الآن نقوم بترتيب البيانات في الجدول (١٢ ـ ٤)

جدول (١٢ ـ ٤)

IQ س ر	الانتاجية ص ر	س ر ص ر	س$^{٢}_{ر}$	ص$^{٢}_{ر}$
١١٠	٥ر٢	٥٧٢	١٢١٠٠	٢٧ر٠٤
١٢٠	٦ر٠	٧٢٠	١٤٤٠٠	٣٦ر٠٠
١٣٠	٦ر٣	٨١٩	١٦٩٠٠	٣٩ر٦٩
١٢٦	٥ر٧	٧١٨ر٢	١٥٨٧٦	٣٢ر٤٩
١٢٢	٤ر٨	٥٨٥ر٦	١٤٨٨٤	٢٣ر٠٤
١٢١	٤ر٢	٥٠٨ر٢	١٤٦٤١	١٧ر٦٤
١٠٣	٣ر٠	٣٠٩	١٠٦٠٩	٩ر٠٠
٩٨	٢ر٩	٢٨٤ر٢	٩٦٠٤	٨ر٤١
٨٠	٢ر٧	٢١٦	٦٤٠٠	٧ر٢٩
٩٧	٣ر٢	٣١٠ر٤	٩٤٠٩	١٠ر٢٤
١١٠٧	٤٤	٥٠٤٢ر٦	١٢٤٨٢٣	٢١٠ر٨٤

من الجدول نجد ما يلي :

$\sum$ س$_{ر}$ = ١١٠٧ ، $\sum$ ص$_{ر}$ = ٤٤ ، $\sum$ س$_{ر}$ ص$_{ر}$ = ٥٠٤٢ر٦

$\sum$ س$^{٢}_{ر}$ = ١٢٤٨٢٣ ، $\sum$ ص$^{٢}_{ر}$ = ٢١٠ر٨٤ ، ن = ١٠ .

وبالتعويض في معادلة معامل الارتباط، نجد قيمة ر

$$\text{ر} = \frac{١٠(٥٠٤٢ر٦) - (١١٠٧)(٤٤)}{\sqrt{١٠(١٢٤٨٢٣) - (١١٠٧)^{٢}}\ \sqrt{١٠(٢١٠ر٨٤) - (٤٤)^{٢}}}$$

$$ر = \frac{٥٠٤٢٦ - ٤٨٧٠٨}{\sqrt{١٢٤٨٢٣٠ - ١٢٥٤٤٩}\ \sqrt{٢١٠٨ر٤ - ١٩٣٦}}$$

$$ر = \frac{١٧١٨}{\sqrt{٢٢٧٨١}\ \sqrt{١٧٢ر٤}}$$

$$ر = \frac{١٧١٨}{(١٥٠ر٩٣)\ (١٣ر١٣)}$$

$$ر = \frac{١٧١٨}{١٩٨٢}$$

$$ر = ٨٦٧ر$$

وهذا يعني ان معامل الارتباط ما بين علامة الذكاء والانتاجية هو حوالي (٨٦٧ر). وبما ان اشارة معامل الارتباط موجبة، فهذا يعني ان العلاقة طردية.

ـ معامل الارتباط للرتب Rank Correlation Coefficoent

لقد بينا في الفصل العاشر ان الاختبارات غير القياسية غالباً ما تكون مناظرة للاختبارات القياسية. ان أحد الاختبارات غير القياسية التي تستخدم لقياس قوة العلاقة ما بين المتغير المستقل والمتغير التابع هو ارتباط الرتب. ويعتبر هذا الاختبار مهم جداً عندما لا نستطيع قياس متغير أو أكثر رقمياً، ولكن من السهل تحديد رتب للصفة المراد دراستها عن المتغير. فعلى سبيل المثال: يريد عالم نفسي تحديد قوة العلاقة ما بين مثابرة الشخص وعلامته في الذكاء (IQ). وبما انه لا يوجد هناك قياس موضوعي للمثابرة، فقد طلب عالم النفس من أحد الأساتذة القيام بترتيب (١٢) طالب بالاعتماد على

وجهة نظره بالنسبة لمثابرة الطلبة. والجدول (١٢ ـ ٥) يبين رتبة المثابرة وعلاقة الذكاء لكل طالب.

الجدول (١٢ ـ ٥)

الشخص	(١) رتبة المثابرة	(٢) IQ	(٣) رتب ألـ(QI)	(٤) $ف_{ر}$	(٥) $ف_{ر}^{٢}$
أ	١	١٢٠	٨	٧	٤٩
ب	٩	١٣٠	٣	-٦	٣٦
ت	٤	١٣٢	٢	-٢	٤
ث	٣	١٢٧	٤	١	١
ج	٨	١٢٦	٥	-٣	٩
ح	٢	١١٠	١١	٩	٨١
خ	١١	٩٨	١٢	١	١
د	١٠	١١٣	١٠	٠	٠
ذ	٦	١٢١	٧	١	١
ر	٥	١٢٣	٦	١	١
ز	٧	١١٩	٩	٢	٤
ع	١٢	١٤٠	١	-١١	١٢١
$\sum ف_{ر}^{٢}$					٣٠٨

ان معامل ارتباط الرتب يقيس قوة العلاقة ما بين مجموعتين من الرتب. بمعنى آخر، ان معامل ارتباط الرتب يقيس العلاقة ما بين رتب المتغير الأول ورتب المتغير الثاني. فعلى سبيل المثال، العمود رقم (٣) في الجدول (١٢ ـ ٥) يبين رتبة كل شخص من الأشخاص بالنسبة لعلامة الذكاء، والعمود رقم (١) يبين رتبة كل شخص من

الأشخاص بالنسبة للمثابرة. فهل هناك ارتباط ما بين هذه الرتب؟ بمعنى آخر، هل الأشخاص الذين لهم رتب عالية بالنسبة لعلامة الذكاء أيضاً لهم رتبة عالية بالنسبة للمثابرة؟ فاذا كان الجواب نعم، فهل هذه العلاقة قوية؟ وللمساعدة في الاجابة على هذه الأسئلة، فانه يمكن استخدام معامل ارتباط سبيرمان للرتب الذي يعرف بالقانون التالي :

$$ر = ١ - \frac{٦ \sum ف_ر^٢}{ن(ن^٢ - ١)}$$

حيث ان ف ر هو الفرق ما بين الرتبتين للمشاهدة (ر)، و (ن) تمثل عدد المشاهدات. لذلك، ان حساب معامل ارتباط الرتب يحتاج الى ايجاد الفرق ما بين رتب علامات الذكاء ورتب المثابرة لكل شخص ثم نعمل على تربيع هذا الفرق ومن ثم نجد مجموع مربعات الفروق في الرتب. وبعد ذلك نقوم بالتعويض في القانون. وكما يظهر في الجدول (١٢ ـ ٥)، نلاحظ ان معامل الارتباط هو

$$ر = ١ - \frac{٦(٣٠٨)}{١٢(١٢^٢ - ١)}$$

$$ر = ١ - \frac{١٨٤٨}{١٧١٦}$$

$$ر = ١ - ١٫٠٧٧$$

$$ر = -٠٫٠٧٧$$

ان قيم معامل الارتباط تترواح ما بين (−١) الى (+١). ان القيمة الموجبة لمعامل ارتباط الرتب تعني ان العلاقة ما بين المتغيرين هي علاقة طردية. والقيمة السالبة لمعامل ارتباط الرتب تعني ان العلاقة ما بين المتغيرين هي علاقة عكسية. أما اذا كانت قيمة معامل ارتباط الرتب = +١ ، فهذا يعني وجود علاقة تامة طردية ما بين المتغيرين. أما اذا

كانت قيمة معامل الارتباط = −١ ، فهذا يعني وجود علاقة تامة وعكسية ما بين المتغيرين. أما اذا كان المتغيرين مستقلين احصائياً، فهذا يعني ان معامل ارتباط الرتب للمتغيرين يساوي صفر.

ـ استنتاج معامل ارتباط المجتمع

Inference Concerning the Population Correlation Coefficient

ان معامل الارتباط الذي تم حسابه في الجزء السابق هو معامل الارتباط العيني، كما ان قيمة هذا المعامل تختلف من عينة الى اخرى. وعلى العكس من ذلك، إن معامل الارتباط للمجتمع (P) (ويقرأ ر و) يتعلق بجميع افراد المجتمع. ان تعريف معامل ارتباط المجتمع (P) هو نفس تعريف معامل ارتباط العينة (ر)، والفرق الوحيد بينهما ان حساب معامل الارتباط العيني يعتمد على عينة من مجتمع الدراسة بينما معامل الارتباط للمجتمع يعتمد على جميع افراد المجتمع. وفي معظم الحالات، يكون اهتمام الشخص الباحث بمعامل ارتباط المجتمع أكثر من معامل ارتباط العينة، لذلك فهم يقومون باستخدام قيمة معامل ارتباط العينة لتقدير قيمة ارتباط المجتمع.

ان اهتمام الباحث غالباً ما يكون في اختبار ما اذا كان معامل الارتباط للمجتمع مساوياً للصفر. فاذا كان متغير ما مستقلاً عن الآخر، فان معامل الارتباط للمجتمع سوف يكون صفراً. فعلى سبيل المثال، اذا كانت انتاجية عامل في وظيفة معينة لا ترتبط على الاطلاق مع عمره، فان معامل الارتباط للمجتمع ما بين الانتاجية والعمر سوف يكون صفراً. لذلك قد يكون اهتمام مدير الانتاج لمؤسسة ما معرفة ما اذا كانت الانتاجية مستقلة فعلاً عن العمر لوظيفة معينة. فاذا تم قياس العلاقة ما بين الانتاجية والعمر لعينة من مجتمع الدراسة، فان باستطاعة مدير الانتاج اختبار الفرضية الأساسية التالية :

ف. : $P = \cdot$

ولانجاز هذا الاختبار، فانه لا بد أولاً من تحديد الفرضية الأساسية والفرضية البديلة. ان الفرضية الأساسية هي

ف. : P = صفر

بينما الفرضية البديلة هي :

ف١ : $P \neq$ صفر

ولاختبار الفرضية، فانه لا بد من حساب

$$t = \frac{ر}{\sqrt{(١ - ر^٢) / (ن - ٢)}}$$

وهذا يعني انه حتى تكون الفرضية الأساسية صحيحة، يجب أن يكون توزيعها هو توزيع (t) بدرجات حرية = ن - ٢ . لذلك فان قانون القرار المستخدم لاختبار الفرضية الأساسية هو: رفض الفرضية الأساسية P = صفر اذا كانت قيمة (t) المحسوبة من العلاقة أعلاه هي أكبر من القيمة المعيارية لـ ($t\frac{\propto}{٢}$) أو أصغر من ($-t\frac{\propto}{٢}$)، وقبول الفرضية الأساسية على العكس من ذلك.

ولتوضيح الطريقة المستخدمة في اختبار الفرضية الأساسية، افرض ان معامل الارتباط ما بين الانتاجية والعمر لعينة مكونة من (١٨) عامل هو (٤٢ر). والمطلوب هو اختبار ما اذا كان معامل الارتباط لمجتمع الدراسة (P) مساوياً للصفر اذا كانت قيمة $\propto$ = ٥٪ .

الحل : لاجراء الاختبار، لا بد من اتباع الخطوات التالية :

١ - تكشيل الفرضية الأساسية والفرضية البديلة.

ف. : P = ٠

ف١ : $P \neq$ ٠

٢ ـ اختيار الاختبار الأمثل.

الاختبار الأمثل هو دائماً اختبار (t) .

٣ ـ تحديد قيمة ∝ .

$$\propto = \text{٥٪}$$

٤ ـ حساب قيمة (t) الاحصائية باستخدام العلاقة التالية :

$$t = \frac{\text{ر}}{\sqrt{(\text{١} - \text{ر}^{\text{٢}}) \,/\, (\text{ن} - \text{٢})}}$$

$$t = \frac{\text{٤٢ر}}{\sqrt{[\text{١} - (\text{٤٢ر})^{\text{٢}}] \,/\, (\text{١٨} - \text{٢})}}$$

$$t = \frac{\text{٤٢ر}}{\sqrt{(\text{٨٢٣٦ر}) \,/\, \text{١٦٨}}}$$

$$t = \frac{\text{٤٢ر}}{\sqrt{\text{٠٥١٥ر}}}$$

$$t = (\text{٠ر٤٢}) \,/\, (\text{٠ر٢٢٦٩})$$

$$t = \text{١ر٨٥١}$$

٥ ـ ايجاد قيمة t المعيارية اذا كانت قيمة

$\propto = \text{٥٪}$ ، د . ح = ١٦ (١٨ – ٢).

$t(\frac{\propto}{\text{٢}}\text{، ١٦}) = \text{٢ر١٢}$ (من الجدول)

٦ ـ القرار : رفض الفرضية الأساسية اذا كانت القيمة المطلقة للقيمة المحسوبة أكبر من القيمة المعيارية أو الجدولية. وعليه فان القرار هو قبول الفرضية الأساسية ورفض الفرضية البديلة لأن القيمة المحسوبة لاختبار (t) هي أقل من القيمة المعيارية لذلك الاختبار. وهذا يعني عدم وجود ارتباط ما بين الانتاجية والعمر.

ـ استنتاج ميل خط الانحدار (ب) . Inference Concerning the Value of B

في تحليل الانحدار، يختلف ميل خط انحدار العينة من عينة الى اخرى. وكما هو الحال بالنسبة لأي عينة احصائية، فان لهذه العينة توزيعاً عينياً. كما أن تقدير الانحراف المعياري للتوزيع العيني هو

$$Sb = Se \div \sqrt{\sum_{ر=1}^{ن} س_ر^٢ - ن \bar{س}^٢}$$

وغالباً ما يطلق هذا المقدار بالخطأ المعياري لـ (ب). ويوجد هناك الكثير من الحالات التي يرغب من خلالها الشخص الباحث من حساب فترة الثقة لميل خط الانحدار للمجتمع (ب) عن طريق قيمة ميل خط الانحدار للعينة (ب). ان حدود فترة الثقة هي :

$$ب \pm (t\tfrac{\propto}{٢})\,(Sb)$$

فاذا كان التوزيع الاحتمالي المشروط للمتغير التابع توزيعاً طبيعياً، فان احتمال ان يقع ميل خط انحدار المجتمع (ب) ضمن حدود هذه الفترة هو (١ − ∝). ولتوضيح استخدام هذه المعادلة، فاننا سوف نقوم بحساب فترة الثقة الـ ٩٥٪ للثابت (ب) لمثال المنشأة التي تريد وصف العلاقة ما بين الانتاج والتكاليف.

الحل : نلاحظ مما سبق ان :

$$ب = ٧٥٢ر \quad،\quad Se = ٣٩٢ر \quad،\quad \sum س_ر^٢ = ٣٤٠$$

$$ن = ٩ \qquad \bar{س} = ٥٥٦ر٥ \quad،\quad t(\tfrac{\propto}{٢}،\ ٧) = ٣٦٥ر٢ \ .$$

لايجاد فترة الثقة، فانه لا بد أولاً من حساب قيمة الخطأ المعياري لـ (ب) والذي يرمز له بالرمز (Sb) .

Sb = Se ÷ √(مجـ سر٢ − ن سَ٢)

Sb = ٣٩٢ر ÷ √(٣٤٠ − ٩(٥ر٥٥٦)٢)

Sb = ٣٩٢ر ÷ √(٣٤٠ − ٢٧٧ر٨٢)

Sb = ٣٩٢ر ÷ √(٦٢ر١٨)

Sb = ٣٩٢ر ÷ ٧ر٨٨٥

Sb = ٠٤٩٧ر

اذن، نستطيع الآن حساب فترة الثقة والتي هي

فترة الثقة = ب ± (t ∝/٢) (Sb)

فترة الثقة = ٧٥٢ر ± (٢ر٣٦٥) (٠٤٩٧ر)

فترة الثقة = ٧٥٢ر ± ١١٨ر

لهذا، فان فترة الثقة الـ ٩٥٪ لتكاليف انتاج (١) طن اضافي هي من (٦٣٤ر) الف دينار الى (٨٧٠ر) الف دينار، أو حدود الفترة هي من ٦٣٤ دينار الى ٨٧٠ دينار.

فبالاضافة الى تقدير قيمة (ب)، فان الباحث غالباً ما يريد اختبار ما اذا كانت قيمة (ب) تساوي صفر. فاذا كانت هذه الفرضية صحيحة، فان الوسط للمتغير التابع هو نفسه وبغض النظر عن قيمة المتغير المستقل. وبالتحديد فان الوسط الحسابي يساوي أ لأن ب س = صفر. بمعنى آخر، اذا كانت هذه الفرضية صحيحة، فانه لا يوجد هناك

علاقة ما بين المتغير التابع والمتغير المستقل.

فاذا كان ميل خط الانحدار للمجتمع (ب) مساوياً للصفر، فانه ليس من الضروري أن يكون ميل خط انحدار العينة (ب) مساوياً للصفر. وعلى العكس من ذلك، ان قيمة (ب) للعينة لا تساوي بالغالب صفراً بسبب التغيرات العشوائية. ان قانون القرار لاختبار الفرضية ب = صفر هو رفض الفرضية الأساسية اذا كانت القيمة المطلقة للقيمة المحسوبة لاختبار (t) أكبر من القيمة المعيارية لاختبار (t) .

ولتوضيح كيفية حساب هذا الاختبار، دعنا نعود لمثال المنشأة اذا أدادت اختبار ما اذا كان ميل خط الانحدار للمجتمع والذي يربط ما بين الانتاج والتكاليف هو صفر، أي أن ب = صفر (ضد الفرضية البديلة ب ≠ صفر)، علماً بأن مستوى الثقة هو ٥٪ .

الحل : من الجزء السابق، نعلم ما يلي :

ب = ٢٥٧ر ، Se = ٢٩٣ر ، $\sum$ س$_{ر}^{2}$ − ن س^{2} = ١٨ر٦٢ ، ن = ٩ .

ولاجراء الاختبار فانه لا بد من اتباع الخطوات التالية :

١ ـ تشكيل الفرضية الأساسية والبديلة.

ف$_{.}$: ب = صفر

ف$_{1}$: ب ≠ صفر

٢ ـ تحديد الاختبار الأمثل.

الاختبار الأمثل هو اختبار t

٣ ـ تحديد مستوى الثقة ($\propto$) .

ان قيمة $\propto$ = ٥٪

٤ ـ حساب القيمة الاحصائية للاختبار (t) من خلال العلاقة التالية :

t = ب ÷ Sb

$$t = \frac{٠٫٧٥٢}{٠٫٣٩٢ \div \sqrt{١٨٫٦٢}}$$

$$t = \frac{٠٫٧٥٢}{٠٫٠٤٩٧}$$

$$t = ١٥٫١$$

٥ ـ ايجاد قيمة (t) المعيارية اذا كانت قيمة $\propto$ = ٥٪ ، د . ح = ٧ (٩ − ٢) . t
($\frac{\propto}{٢}$ ، ٧) = ٢٫٣٦٥

٦ ـ القرار : بما ان القيمة المحسوبة لاختبار (t) اكبر من القيمة المعيارية، فهذا يعني رفض الفرضية الأساسية وقبول الفرضية البديلة. وهذا يعني ان ميل خط الانحدار للمجتمع هو قيمة موجبة وليس صفراً.

مثال (٦) : بالاعتماد على البيانات الواردة في المثال رقم (١)، اختبر الفرضية التي تعتبر ان ميل خط الانحدار الصحيح والذي يصف العلاقة ما بين الدخل والادخار مساوية للصفر، اذا كانت قيمة $\propto$ = ٥٪ .

الحل : من نتائج المثال رقم (١) والمثال رقم (٢)، نلاحظ ان :

ب = ١٫٠١٧ ، Se = ٠٫٥٣ ، $\sum$ س$_{ر}^{٢}$ − ن س̄^{٢} = ٦٠ ، ن = ٩ .

وعليه لاجراء الاختبار فانه لا بد من اتباع الخطوات التالية :

١ ـ تشكيل الفرضية الأساسية والفرضية البديلة.

ف$_{٠}$: ب = صفر

ف$_{١}$: ب ≠ صفر

٢ ـ تحديد الاختبار الأمثل.

الاختبار الأمثل هو اختبار (t)

٣ ـ تحديد مستوى الثقة (∝) .

قيمة ∝ = ٥٪

٤ ـ ايجاد القيمة الاحصائية للاختبار (t) .

$t = \text{ب} \div S_b$

$$t = \frac{\text{١٠١٧ر}}{\text{٠٥٣ر} \div \sqrt{\text{٦٠}}}$$

$$t = \frac{\text{١٠١٧ر}}{\text{٠٠٦٨ر}}$$

$t = \text{١٤ر٩٥}$

٥ ـ ايجاد قيمة t المعيارية اذا كانت

∝ = ٥٪ ، د . ح = ٧ ، $t(\frac{\propto}{\text{٢}} ، \text{٧})$ = ٢ر٣٦٥ .

٦ ـ القرار : رفض الفرضية الأساسية وقبول الفرضية البديلة لأن القيمة المحسوبة لاختبار (t) أكبر من القيمة المعيارية. وهذا يعني ان ميل خط الانحدار لمجتمع الدراسة هو قيمة موجبة وليس صفراً.

أسئلة

س١ : اذا كانت نسبة الزيادة في الانفاق على البحث والتطوير لخمسة عشر شركة كيماوية أردنية (متغير مستقل) ونسبة الزيادة في الانفاق على البحث والتطوير لخمسة عشر شركة كيماوية غير أردنية (متغير تابع) مزودة أدناه

الشركة	س	ص
١	٥ر٠٠	٠ر٠٠
٢	٨ر٩٠	٠ر٤٣
٣	٧ر١٥	٠ر٠٠
٤	٦ر١٠	٠ر٢٤
٥	٧ر٧٠	٠ر٠٠
٦	٨ر٢٠	٠ر٩١
٧	١ر٠١	٠ر٠٠
٨	٠ر٦١	٠ر٠٣
٩	٠ر٧٢	٠ر٠١
١٠	٠ر٦٨	٠ر٠٠
١١	١ر١٤	٠ر٠٠
١٢	١ر١٨	٠ر٠١
١٣	٠ر٧٣	٠ر٠٠
١٤	٠ر٨٧	٠ر٢٠
١٥	١ر٤٧	٠ر٠٠

أوجد ما يلي :

أ - ارسم لوحة الانتشار لهذه البيانات .

ب - احسب معادلة خط الانحدار للعينة.

جـ - ما هي نسبة الزيادة المتوقعة في الانفاق على البحث والتطوير خارج الأردن اذا أنفقت شركة كيماوية أردنية ٦٪ زيادة على البحث والتطوير.

د ـ هل العلاقة ما بين المتغير المستقل والتابع علاقة سببية، لماذا ؟

س٢ : حصال عالم نفس على علامات الذكاء (IQ) لأربعة عشر شخصاً، وطلب من كل شخص تقديم امتحان معين، فكانت النتائج كما يلي :

IQ	علامة الامتحان
٩٩	٥٤
١١٠	٧٠
١٤١	١٠٠
١٠٨	٦١
١٢٣	٨٣
١٢٩	٨٦
١٣٢	٩٨
٨٨	٥١
١٠١	٦٤
١٠٥	٦٨
٩٧	٥٢
٩٦	٥٥
١١٨	٧١
١١٣	٧٦

أوجد ما يلي :

أ - ارسم لوحة الانتشار لهذه البيانات.

ب - احسب معادلة خط الانحدار للعينة.

جـ - ما هي علامة الامتحان المتوقعة لشخص علامة ذكاه هي ١٣٠.

س٣ : اذا كانت البيانات المزودة تمثل العلاقة ما بين سعر الكتاب وعدد الصفحات، فأوجد ما يلي :

أ - ارسم لوحة الانتشار لهذه البيانات.

ب - احسب معادلة خط الانحدار للعينة.

جـ - احسب الخطأ المعياري للتقدير.

د - احسب فترة الثقة الـ ٩٥٪ لمتوسط السعر المشروط، اذا كان عدد صفحات الكتاب (٦٠٠) صفحة.

هـ - احسب فترة الثقة الـ ٩٥٪ لسعر الكتاب، اذا كان عدد صفحاته هي (٦٠٠) صفحة.

السعر (بالدينار)	عدد صفحات الكتاب
٩	٥٢٠
١٣	٦٨٠
١٤	٧٤٠
٢	٢٠٠
٣	٤٠٠
١٥	٨٠٠
١٣	٧٥٠
٧	٥٠٠
٣	٣٠٠
٣	٣٥٠

س٤ : اذا كان مقدار الانفاق على الطعام والدخل السنوي لتسع عائلات هي كما يلي :

الانفاق على الطعام (١٠٠٠)	الـدخـل (١٠٠٠)
٤	٢٠
٦	٤٠
٣	١١
٥	٣٠
٢	٩
٢	١٢
٣	١٥
٣	٢١
٨	٥٠

أوجد ما يلي :

أ ـ احسب معادلة خط الانحدار للعينة.

ب ـ احسب الخطأ المعياري للتقدير.

جـ ـ احسب فترة الثقة الـ ٩٠٪ لمتوسط الانفاق المشروط على الطعام لعائلة دخلها السنوي هو (٢٠) الف دينار.

د ـ احسب فترة الثقة الـ ٩٠٪ لانفاق عائلة على الطعام اذا كان دخلها السنوي هو (٢٠) الف دينار.

س٥ : استخدم البيانات الموجودة في السؤال رقم (٢) في ايجاد ما يلي :

أ ـ أوجد الخطأ المعياري للتقدير.

ب ـ احسب فترة الثقة الـ ٩٩٪ لعلامة شخص في الاختبار اذا كانت علامة الذكاء له هي (١٣٠) .

س٦ : بناءاً على البيانات المنشورة في مراقبة النوعية الصناعية، كانت العلاقة ما بين المادة المنظمة (ص) والتركيز (س) لعينة مكونة من ثمانية مواد منظمة كما يلي :

ص	س
٣٧	١٠
٤٢	٢٠
٤٦	٣٠
٤٨	٤٠
٥٣	١٠
٦٢	١٠
٧٩	٣٠
٨٤	٤٠

أوجد ما يلي :

أ ـ احسب معامل الارتباط.

ب ـ اختبر ما اذا كان معامل الارتباط للمجتمع يساوي صفر (اعتبر $\propto$ = ٠٥ و).

س٧ : اذا كانت العلاقة ما بين المعدل الجامعي ومعدل الثانوية العامة لعينة عشوائية مكونة من (١٠) طلاب هي كما يلي :

المعدل الجامعي	معدل الثانوية العامة
٦٥	٧٦
٥٠	٦٧
٧٨	٧٧
٨٨	٨٢
٩٥	٩٩
٥٥	٧٥
٦٨	٨٤
٧٠	٨٧
٧٥	٩٠
٨٤	٩٩

أوجد ما يلي :

أ ـ احسب معامل الارتباط.

ب ـ اختبر ما اذا كان معامل الارتباط للمجتمع يساوي صفر (اعتبر ان $\propto$ = ٠١ ر) .

س٨ : بالاعتماد على البيانات الواردة في السؤال رقم (١)، أوجد معامل التحديد، فسر الاجابة.

س٩ : بالاعتماد على البيانات الواردة في السؤال رقم (١)، أوجد ما يلي :

أ - معامل التحديد، فسر اجابتك.

ب - اختبر ما اذا كان ميل خط الانحدار للمجتمع يساوي صفر (اعتبر $\propto$ = ٥٪) .

جـ - احسب فترة الثقة ٩٠٪ لميل خط الانحدار الصحيح.

س١٠ : بالاعتماد على البيانات الواردة في السؤال رقم (٢)، أوجد ما يلي :

أ - معامل التحديد، فسر اجابتك.

ب - اختبر ما اذا كان ميل خط الانحدار للمجتمع يساوي صفر (اعتبر $\propto$ = ١ر٠) .

جـ - احسب فترة الثقة الـ ٩٠٪ لميل خط الانحدار الصحيح.

د - اختبر ما اذا كان معامل الارتباط للمجتمع يساوي صفر (اعتبر $\propto$ = ١ر).

س١١ : اذا كان معامل الارتباط للعينة هو (٣و) وكان حجم العينة ن = ٢٠، اختبر ما اذا كان معامل الارتباط للمجتمع يساوي صفر، اذا كانت قيمة $\propto$ = ٥٪ .

الفصل الثالث عشر

الانحدار المتعدد والارتباط

Multiple Regression and Correlation

الانحدار المتعدد والارتباط

Multiple Regression and Correlation

- مقدمة

لقد قمنا بالفصل السابق بشرح أساليب الانحدار والارتباط في الحالات التي يكون فيها متغير مستقل واحد فقط. ان التطبيق العملي لتحليل الانحدار والارتباط غالباً ما يتضمن متغيرين أو أكثر من المتغيرات المستقلة. وفي هذا الفصل سوف نقوم بتطبيق تحليل الانحدار والارتباط في الحالات التي يكون فيها أكثر من متغير مستقل.

- تحليل الانحدار : طبيعته وأغراضه

Multiple Regression Nature and Purpose

اذا كان الانحدار البسيط يتضمن متغير مستقل واحد، فان الانحدار المتعدد يتضمن متغيرين أو أكثر من المتغيرات المستقلة. وبشكل عام، يوجد هناك سببان رئيسيان يوجبان استخدام تحليل الانحدار المتعدد بدلاً من الانحدار البسيط. وهذان السببان هما :

١ - ان تنبؤ المتغير التابع غالباً ما يكون أكثر دقة اذا تم استخدام أكثر من متغير مستقل. ففي حالة المنشأة التي تم دراستها في الفصل السابق، قد يشعر قسم البحث في المنشأة بوجود عوامل أخرى مهمة لها تاثير على تكاليف المنشأة بالاضافة الى كمية الانتاج. فعلى سبيل المثال، إنه من المتوقع أن تتناقص تكاليف المنشأة اذا كانت خبرة المدراء عالية. لذلك، اذا تم افتراض ثبات كمية الانتاج، فان القيمة المتوقعة لتكاليف

المنشأة قد تكون اقتران خطي بالنسبة لعدد سنوات الخبرة عند مدراء. بمعنى آخر، انه من المعقول والمنطقي جداً افتراض ما يلي :

$$ت(ص_ر) = أ + ب_١ س_{١ر} + ب_٢ س_{٢ر} \quad .... (١٣-١)$$

حيث ان :

$ص_ر$ = تكلفة الشهر (ر) بآلاف الدنانير.

$س_{١ر}$ = كمية الانتاج لذلك الشهر.

$س_{٢ر}$ = عدد سنوات الخبرة للمدراء المسؤولين في المنشأة في ذلك الشهر.

بالطبع، اذا كانت زيادة سنوات الخبرة عند المدراء تؤدي الى تقليل التكاليف، فان قيمة الثابت ($ب_٢$) يجب ان تكون سالبة.

اعتماداً على المعادلة (١٣ - ١) ، ان القيمة المتوقعة للتكاليف في شهر معين تعتمد على كمية الانتاج لذلك الشهر وعلى سنوات الخبرة لمدراء المنشأة في ذلك الشهر. ان المعادلة تقول، اذا زادت كمية الانتاج بمقدار (١) طن، فان القيمة المتوقعة للتكاليف سوف تزداد بمقدار قيمة ($ب_١$) الف دينار. واذا زادت سنوات الخبرة لمدراء المنشأة سنة واحدة، فان قيمة التكاليف سوف تزداد بمقدار قيمة ($ب_٢$) الف دينار. وبما انه من المفروض ان تكون قيمة ($ب_١$) قيمة سالبة، فهذا يعني انه من المفروض ان تتناقص تكاليف المنشأة بقيمة ($ب_٢$) اذا زادت سنوات خبرة مدراء المنشأة بمقدار سنة واحدة. فاذا كان صحيح ان تكلفة الانتاج تعتمد على كمية الانتاج وعدد سنوات الخبرة، فهذا يعني ان قسم البحث او قسم الانتاج في المنشأة قادراً على توّقع تكاليف المنشأة بصورة افضل عن طريق استخدام تحليل الانحدار المتعدد. بمعنى آخر، ان استخدام معادلة خط الانحدار المتعدد التي تربط المتغير التابع (تكاليف المنشأة) بكلا المتغيرين المستقلين (كمية الانتاج وعدد سنوات الخبرة) يكون افضل من استخدام معادلة خط الانحدار البسيط التي تم شرحها في الفصل السابق والتي تربط المتغير التابع

(تكاليف المنشأة) مع متغير مستقل واحد فقط (كمية الانتاج) .

٢ ـ ان السبب الثاني لاستخدام خط الانحدار المتعدد بدلاً من خط الانحدار البسيط هو ان المتغير التابع غالباً ما يعتمد على أكثر من مستقل. بمعنى آخر، ان تحليل الانحدار البسيط الذي يعتمد على متغير مستقل واحد قد يؤدي الى تحيّز في تقدير تاثير المتغير المستقل على المتغير التابع. فعلى سبيل المثال، افرض ان تكاليف المنشأة وبالاعتماد على المعادلة (١٣ ـ ١) تعتمد على كمية الانتاج وسنوات الخبرة، فان تقدير التكاليف بناءاً على كمية الانتاج وباستخدام خط الانحدار البسيط قد يؤدي الى تحيّز في تقدير قيمة ($ب_1$) والتي تقيس الزيادة بالتكاليف بناءاً على انتاج (١) طن اضافي. ولفهم سبب ظهور هذا التحيز، افرض ان المنشأة أرادت زيادة انتاجها في الأشهر التي زادت فيها خبرة مدراءها عن الأشهر التي كانت فيها خبرة المدراء قليلة نسبياً. فاذا كانت هذه هي الحالة، فان تكاليف الأشهر التي كانت غالباً فيها كمية الانتاج قليلة كانت عالية بسبب قلة خبرة المدراء، وتكاليف الأشهر التي كانت فيها كمية الانتاج عالية كانت قليلة بسبب زيادة خبرة المدراء. لهذا، فان استخدام تحليل الانحدار البسيط لوصف العلاقة ما بين التكاليف والانتاج فقط سوف يؤدي الى تحيّز في تقدير قيمة ($ب_1$) .

اذا كان المتغير التابع هو اقتران لأكثر من متغير مستقل، فان العلاقة الملاحظة ما بين المتغير التابع وأي متغير من المتغيرات المستقلة قد تكون مضللة لأن العلاقة الملاحظة قد تعكس التغير في المتغيرات المستقلة الأخرى، وبما ان هذه المتغيرات المستقلة الأخرى غير مراقبة كلياً، فان هذا قد يؤدي الى إظهار تاثير المتغير المستقل المراقب على المتغير التابع بصورة غير صحيحة. ولتقدير التأثير الصحيح للمتغير المستقل على المتغير التابع، فانه لا بد من تضمين جميع المتغيرات المستقلة في تحليل الانحدار. بمعنى آخر، لا بد من استخدام نموذج خط الانحدار المتعدد.

ـ نموذج خط الانحدار المتعدد The Multiple Regression Model

ان النموذج الأساسي لخط الانحدار البسيط وكما تبين من الفصل السابق هو كما يلي :

ص$_{ر}$ = أ + ب س + خ$_{ر}$ .... (١٣-٢)

حيث ان :

ص$_{ر}$ هي القيمة الملاحظة (ر) من قيم المتغير التابع.

س$_{ر}$ هي القيمة الملاحظة (ر) من قيم المتغير المستقل.

خ$_{ر}$ هو المتغير ذو التوزيع العشوائي الطبيعي والذي وسطه الحسابي يساوي صفراً وانحرافه المعياري هو (σe). بمعنى آخر، تمثل قيمة خ ر مقدار الخطأ، أو مقدار القيمة العشوائية المضافة الى معادلة خط الانحدار أو المطروحة اذا كانت قيمة خ$_{ر}$ قيمة سالبة.

كما اننا نفترض ان الوسط المشروط للمتغير التابع هو اقتران خطي للمتغير ص$_{ر}$ - وبالتحديد أ + ب س. كما نفترض ان قيم مقدار الخط (خ$_{ر}$) تكون مستقلة احصائياً.

ان نموذج الانحدار الخطي المتعدد هو نفس نموذج الانحدار الخطي البسيط الذي تم الاشارة له أعلاه. ان الفرق الوحيد ما بين هذين النموذجين هو ان الوسط المشروط للمتغير التابع هو اقتران خطي لأكثر من متغير مستقل. فاذا كان هناك متغيرين مستقلين هما س١ و س٢ ، فان النموذج الخطي المتعدد هو :

ص$_{ر}$ = أ + ب$_{١}$ س$_{١ر}$ + ب$_{٢}$ س$_{٢ر}$ + ح$_{ر}$

حيث أن خ$_{ر}$ هو مقدار الخطأ. وكما هو الحال بالنسبة للتحليل الانحدار البسيط، فان متوسط القيمة المتوقعة لمقدار الخطأ هو صفر، كما ان مقدار الخطأ موزع طبيعياً والانحراف المعياري له هو نفسه وبغض النظر عن قيم كل من س$_{١ر}$ و س$_{٢ر}$. كما نفترض ان قيم مقدار الخطأ تكون مستقلة احصائياً. ولكن على العكس من حالة الانحدار البسيط، ان الوسط المشروط للمتغير التابع (ص$_{ر}$) هو اقتران خطي لكل من

المتغيرين المستقلين $س_{1ر}$ و $س_{2ر}$. وبالتحديد ، فان الوسط المشروط للمتغير التابع ($ص_ر$) هو :

$$أ + ب_1 س_{1ر} + ب_2 س_{2ر}$$

ـ طريقة أقل المربعات في تقدير معاملات الانحدار المتعدد

Least-Squares Estimates of the Regression Cofficient

ان أول خطوة في تحليل الانحدار المتعدد هو تحديد المتغيرات المستقلة، ومن ثم تحديد الشكل الرياضي للمعادلات ذات العلاقة بمتوسط القيمة للمتغير التابع بالنسبة للمتغيرات المستقلة. وبالنسبة لحالة المنشأة التي تريد وصف العلاقة ما بين التكلفة والانتاج وسنوات الخبرة فقد تم عمل هذه الخطوة في المعادلة (١٣ - ١) والتي تشير الى ان المتغيرات المستقلة هي كمية الانتاج وعدد سنوات الخبرة. ان العلاقة ما بين متوسط القيمة للمتغير التابع (التكلفة) وهذه المتغيرات المستقلة هي علاقة خطية. فبعد القيام بأول خطوة ، فاننا بعد ذلك نقوم بتقدير قيم الثوابت المجهولة أ ، $ب_1$ ، $ب_2$ في خط الانحدار الصحيح (خط انحدار المجتمع). وكما كان عليه الحال في تحليل الانحدار البسيط، فان تقدير هذه القيم قد تم بناءاً على ايجاد قيمة كل ثابت بحيث يؤدي الى تقليل مجموع مربع الانحرافات لقيم المتغير المستقل الملاحظة عن قيم المتغير التابع المتوقعة باستخدام تحليل الانحدار.

ولفهم طبيعة أقل المربعات في تقدير قيم الثوابت أ ، $ب_1$ ، $ب_2$ بدقة أكبر، افرض ان أ هي قيمة تقديرية لـ أ ، $ب_1$ هي قيمة تقديرية لـ ($ب_1$) ، و $ب_2$ هي قيمة تقديرية لـ ($ب_2$) ، فان قيمة المتغير التابع ($\hat{ص}_ر$) التقديرية المتوقعة عن طريق تحليل الانحدار هي :

$$\hat{ص}_ر = أ + ب_1 س_{1ر} + ب_2 س_{2ر}$$

كما ان انحراف هذه القيمة المتوقعة عن القيمة الحقيقية للمتغير التابع هي :

$$ص_ر - \hat{ص}_ر = ص_ر - أ - ب_١ س_{١ر} - ب_٢ س_{٢ر}$$

وكما هو الحال بالنسبة لتحليل الانحدار البسيط، فانه يمكن قياس درجة تمثيل خط الانحدار التقديري للبيانات عن طريق تربيع هذه الانحرافات :

$$\sum (ص_ر - \hat{ص})^٢ = \sum (ص_ر - أ - ب_١ س_{١ر} - ب_٢ س_{٢ر})^٢ \quad ..... (١٣-٣)$$

حيث ان ن هي مجموع المشاهدات في العينة. فكلما زاد مربع مجموع هذا المقدار، فكلما قل تمثيل معادلة خط الانحدار المتعدد للعينة للبيانات. وكلما قل مجموع مربع هذا المقدار، كلما زاد تمثيل خط الانحدار المتعدد للعينة للبيانات. لذلك، فانه من المنطقي أن نختار قيم الثوابت أ، ب١ و ب٢ والذي يقلل من قيمة المقدار (١٣ - ٣). ان هذه القيم هي القيم المقدرة بطريقة اقل المربعات وكما هو الحال بالنسبة للانحدار البسيط.

ان قيم الثوابت أ ، $ب_١$ ، $ب_٢$ والتي تؤدي الى تقليل مجموع مربع الانحرافات في المعادلة (١٣ - ٣) يجب ان تحقق المعادلات التالية :

$$\sum_{ر=١}^{ن} ص_ر = ن أ + ب_١ \sum_{ر=١}^{ن} س_{١ر} + ب_٢ \sum_{ر=١}^{ن} س_{٢ر}$$

$$\sum_{ر=١}^{ن} س_{١ر} ص_ر = أ \sum_{ر=١}^{ن} س_{١ر} + ب_١ \sum_{ر=١}^{ن} س_{١ر}^٢ + ب_٢ \sum_{ر=١}^{ن} س_{١ر} س_{٢ر}$$

$$\sum_{ر=١}^{ن} س_{٢ر} ص_ر = أ \sum س_{٢ر} + ب_١ \sum س_{١ر} س_{٢ر} + ب_٢ \sum س_{٢ر}^٢$$

.... (١٣ - ٤)

ان حل هذه المعادلة أعلاه بالنسبة للثوابت أ، ب١ ، ب٢ يؤدي الى النتائج التالية :

$$\text{ب}_1 = \frac{\sum(\text{س}_{2\text{ر}}-\bar{\text{س}}_2)^2 \sum(\text{س}_{1\text{ر}}-\bar{\text{س}}_1)(\text{ص}_\text{ر}-\bar{\text{ص}}) - \sum(\text{س}_{1\text{ر}}-\bar{\text{س}}_1)(\text{س}_{2\text{ر}}-\bar{\text{س}}_2) \sum(\text{س}_{2\text{ر}}-\bar{\text{س}}_2)(\text{ص}_\text{ر}-\bar{\text{ص}})}{\sum(\text{س}_{1\text{ر}}-\bar{\text{س}}_1)^2 \sum(\text{س}_{2\text{ر}}-\bar{\text{س}}_2)^2 - [\sum(\text{س}_{1\text{ر}}-\bar{\text{س}}_1)(\text{س}_{2\text{ر}}-\bar{\text{س}}_2)]^2}$$

$$\text{ب}_2 = \frac{\sum(\text{س}_{1\text{ر}}-\bar{\text{س}}_1)^2 \sum(\text{س}_{2\text{ر}}-\bar{\text{س}}_2)(\text{ص}_\text{ر}-\bar{\text{ص}}) - \sum(\text{س}_{1\text{ر}}-\bar{\text{س}}_1)(\text{س}_{2\text{ر}}-\bar{\text{س}}_2) \sum(\text{س}_{1\text{ر}}-\bar{\text{س}}_1)(\text{ص}_\text{ر}-\bar{\text{ص}})}{\sum(\text{س}_{1\text{ر}}-\bar{\text{س}}_1)^2 \sum(\text{س}_{2\text{ر}}-\bar{\text{س}}_2)^2 - [\sum(\text{س}_{1\text{ر}}-\bar{\text{س}}_1)(\text{س}_{2\text{ر}}-\bar{\text{س}}_2)]^2}$$

$$\text{أ} = \bar{\text{ص}} - \text{ب}_1\,\bar{\text{س}}_1 - \text{ب}_2\,\bar{\text{س}}_2$$

(١٣ – ٥) ....

ولتسهيل العمليات الحسابية، فانه لا بد من مراعاة ما يلي :

$$\sum_{\text{ر}=1}^{\text{ن}}(\text{س}_{1\text{ر}}-\bar{\text{س}}_1)^2 = \sum_{\text{ر}=1}^{\text{ن}}\text{س}^2_{1\text{ر}} - \frac{(\sum_{\text{ر}=1}^{\text{ن}}\text{س}_{1\text{ر}})^2}{\text{ن}}$$

$$\sum_{\text{ر}=1}^{\text{ن}}(\text{س}_{2\text{ر}}-\bar{\text{س}}_2)^2 = \sum_{\text{ر}=1}^{\text{ن}}\text{س}^2_{2\text{ر}} - \frac{(\sum_{\text{ر}=1}^{\text{ن}}\text{س}_{2\text{ر}})^2}{\text{ن}}$$

$$\sum_{\text{ر}=1}^{\text{ن}}(\text{س}_{1\text{ر}}-\bar{\text{س}}_1)(\text{ص}_\text{ر}-\bar{\text{ص}}) = \sum_{\text{ر}=1}^{\text{ن}}\text{س}_{1\text{ر}}\,\text{ص}_\text{ر} - \frac{(\sum_{\text{ر}=1}^{\text{ن}}\text{س}_{1\text{ر}})(\sum_{\text{ر}=1}^{\text{ن}}\text{ص}_\text{ر})}{\text{ن}}$$

$$\sum_{\text{ر}=1}^{\text{ن}}(\text{س}_{2\text{ر}}-\bar{\text{س}}_2)(\text{ص}_\text{ر}-\bar{\text{ص}}) = \sum_{\text{ر}=1}^{\text{ن}}\text{س}_{2\text{ر}}\,\text{ص}_\text{ر} - \frac{(\sum_{\text{ر}=1}^{\text{ن}}\text{س}_{2\text{ر}})(\sum_{\text{ر}=1}^{\text{ن}}\text{ص}_\text{ر})}{\text{ن}}$$

$$\sum_{\text{ر}=1}^{\text{ن}}(\text{س}_{1\text{ر}}-\bar{\text{س}}_1)(\text{س}_{2\text{ر}}-\bar{\text{س}}_2) = \sum_{\text{ر}=1}^{\text{ن}}\text{س}_{1\text{ر}}\,\text{س}_{2\text{ر}} - \frac{(\sum_{\text{ر}=1}^{\text{ن}}\text{س}_{1\text{ر}})(\sum_{\text{ر}=1}^{\text{ن}}\text{س}_{2\text{ر}})}{\text{ن}}$$

ولتوضيح العملية الحسابية لكيفية تقدير الثوابت أ ، $ب_1$ ، $ب_2$ باستخدام طريقة أقل المربعات، فاننا سوف نقوم بحل المثال التالي. وعلى الرغم من امكانية حل هذه النوع من الأمثلة باستخدام أجهزة الحاسوب، الا انه من المفيد جداً عمل هذا النوع من الحسابات بالطريقة اليدوية على الأقل لمثال واحد.

مثال (١) : اذا أرادت المنشأة التي تم دراستها في الفصل السابق وصف العلاقة ليس فقط ما بين كمية الانتاج والتكاليف. انما تريد وصف العلاقة ما بين تكلفة الانتاج الشهرية وكمية الانتاج الشهري وعدد سنوات الخبرة عند المدراء. فاذا تم الحصول على بيانات بالنسبة للتكلفة وكمية الانتاج وعدد سنوات الخبرة لتسعة اشهر وكما هي في الجدول (١٣ - ١). فأوجد ما يلي :

(١) معادلة خط الانحدار المتعدد عن طريق طريقة اقل المربعات لتقدير قيمة كل من أ، ب١، ب٢.

(٢) قارن ما بين القيمة التقديرية لـ ب التي تم الحصول عليها باستخدام تحليل الانحدار البسيط وقيمة ب١، هل هناك فرق بينهما، ولماذا ؟

(٣) اي القيمتين التقديريتين برأيك افضل، ولماذ ؟

جدول (١٣ - ١)

التكلفة الشهرية وكمية الانتاج، والخبرة للمنشأة خلال تسعة أشهر

كمية الانتاج (بالطن)	تكاليف الانتاج (بالأف)	سنوات خبرة المدراء
١	٢	١
٢	٣	٠
٤	٤	٥
٨	٧	٨
٦	٦	٤
٥	٥	٣
٨	٨	٢
٩	٨	٧
٧	٦	٦

الحل : لحل هذا المثال، فانه لا بد من ترتيب البيانات في جدول (١٣ ـ ٢)

جدول (١٣ ـ ٢)

الانتاج $س_{١ر}$	التكاليف $ص_{ر}$	الخبرة $س_{٢ر}$	$س^{٢}_{١ر}$	$س^{٢}_{٢ر}$	$س_{١ر}\ ص_{ر}$	$س_{٢ر}\ ص_{ر}$	$س_{١ر}\ س_{٢ر}$	$ص^{٢}_{ر}$
١	٢	١	١	١	٢	٢	١	٤
٢	٣	٠	٤	٠	٦	٠	٠	٩
٤	٤	٥	١٦	٢٥	١٦	٢٠	٢٠	١٦
٨	٧	٨	٦٤	٦٤	٦٤	٥٦	٦٤	٤٩
٦	٦	٤	٣٦	١٦	٣٦	٢٤	٢٤	٣٦
٥	٥	٣	٢٥	٩	٢٥	١٥	١٥	٢٥
٨	٨	٢	٦٤	٤	٦٤	١٦	١٦	٦٤
٩	٨	٧	٨١	٤٩	٨١	٥٦	٦٣	٦٤
٧	٦	٦	٤٩	٣٦	٤٩	٣٦	٤٢	٣٦
٥٠	٤٩	٣٦	٣٤٠	٢٠٤	٣١٩	٢٢٥	٢٤٥	٣٠٣

حيث ان : $س_{١ر}$ تمثل الانتاج بالطن.

$س_{٢ر}$ تمثل عدد سنوات الخبرة.

$ص_{ر}$ تمثل تكاليف الانتاج بالطن.

فبالاعتماد على البيانات الواردة في الجدول (١٣ ـ ٢)، نجد ان

$\sum س_{١ر} = ٥٠$ ، $\sum س_{٢ر} = ٣٦$ ، $\sum ص_{ر} = ٤٩$

$$\sum \text{س}^{2}_{1ر} = \text{٣٤٠} \quad ، \quad \sum \text{س}^{2}_{2ر} = \text{٢٠٤} \quad ، \quad \sum \text{ص}^{2}_{ر} = \text{٣٠٣}$$

$$\sum \text{س}_{1ر}\,\text{س}_{2ر} = \text{٢٤٥} \; ، \; \sum \text{س}_{1ر}\,\text{ص}_{ر} = \text{٣١٩} \; ، \; \sum \text{س}_{2ر}\,\text{ص}_{ر} = \text{٢٢٥} \; .$$

ان تعويض هذه القيم في المعادلات التي تسهل العلمليات الحسابية، نجد ما يلي :

١- $$\sum (\text{س}_{1ر} - \bar{\text{س}}_{1})^{2} = \text{٣٤٠} - \frac{(\text{٥٠})^{2}}{\text{٩}}$$

$$= \text{٣٤٠} - \text{٢٥٠٠} / \text{٩}$$

$$= \text{٣٤٠} - \text{٢٧٧٫٧٨}$$

$$= \text{٦٢٫٢٢}$$

٢- $$\sum (\text{س}_{2ر} - \bar{\text{س}}_{2}) = \text{٢٤٠} - \frac{(\text{٣٦})^{2}}{\text{٩}}$$

$$= \text{٢٤٠} - \text{١٢٩٦} / \text{٩}$$

$$= \text{٢٤٠} - \text{١٤٤}$$

$$= \text{٦٠}$$

٣- $$\sum (\text{س}_{1ر} - \bar{\text{س}}_{1})(\text{ص}_{ر} - \bar{\text{ص}}) = \text{٣١٩} - \frac{(\text{٥٠})(\text{٤٩})}{\text{٩}}$$

$$= \text{٣١٩} - \text{٢٤٥٠} / \text{٩}$$

$$= \text{٣١٩} - \text{٢٧٢٫٢٢}$$

$$= \text{٤٦٫٧٨}$$

٤- $$\sum (\text{س}_{2ر} - \bar{\text{س}}_{2})(\text{ص}_{ر} - \bar{\text{ص}}) = \text{٢٢٥} - \frac{(\text{٣٦})(\text{٤٩})}{\text{٩}}$$

= ٢٢٥ − ١٧٦٤ /٩

= ٢٢٥ − ١٩٦

= ٢٩

٥ - Σ(س١ر − سَ١)(س٢ر − سَ٢) = ٢٤٥ − (٥٠)(٣٦) / ٩

= ٢٤٥ − ١٨٠٠ /٩

= ٢٤٥ − ٢٠٠

= ٤٥

أما الآن فاننا سوف نقوم بتعويض نتائج المعادلات التسهيلة في المعادلات (١٣ - ٥)، ونحصل على ما يلي :

ب١ = [(٦٠)(٤٦ر٧٨) − (٤٥)(٢٩)] / [(٦٢ر٢٢)(٦٠) − (٤٥)^٢]

ب١ = [٢٨٠٦ر٨ − ١٣٠٥] / [٣٧٣٣ر٢ − ٢٠٢٥]

ب١ = ١٥٠١ر٨ / ١٧٠٨ر٢

ب١ = ٨٧٩ر

وهذا يعني انه كلما زاد الانتاج بمقدار طن واحد، فانه سوف يتبع ذلك زيادة بالتكاليف قدرها (٨٧٩) دينار.

$$\text{ب}_{٢} = \frac{(\text{٦٢ر٢٢})(٢٩) - (٤٥)(\text{٤٦ر٧٨})}{(\text{٦٢ر٢٢})(٦٠) - (٤٥)^{٢}}$$

$$\text{ب}_{٢} = \frac{\text{١٨٠٤ر٣٨} - \text{٢١٠٥ر١}}{\text{٣٧٣٣ر٢} - ٢٠٢٥}$$

$$\text{ب}_{٢} = \frac{-\text{٣٠٠ر٧٢}}{\text{١٧٠٨ر٢}}$$

$$\text{ب}_{٢} = -\text{ر١٧٦}$$

وهذا يعني انه كلما زادت الخبرة بمقدار سنة واحدة، فانه سوف يتبع ذلك نقصان بالتكاليف قدرها (١٧٦) دينار.

$$\text{أ} = \frac{٤٩}{٩} - (\text{ر٨٧٩})\left(\frac{٥٠}{٩}\right) - (-\text{ر١٧٦})\left(\frac{٣٦}{٩}\right)$$

$$= \text{٥ر٤٤٤} - \text{٤ر٨٨٣} + \text{ر٧٠٤}$$

$$= \text{١ر٣٦٥}$$

وهذا يعني انه لو كانت كمية الانتاج صفر طن والخبرة صفر سنة، فان التكلفة الثابتة سوف تكون (١ر٣٦٥) دينار.

وعليه فان معادلة خط الانحدار المتعدد التقديرية هي :

$$\hat{\text{ص}}_{\text{ر}} = \text{١ر٣٦٥} + \text{ر٨٧٩}\ \text{س}_{\text{١ر}} - \text{ر١٧٦}\ \text{س}_{\text{٢ر}}$$

ان القيمة التقديرية للثابت ب١ هي (٨٧٩ر)، مقارنة مع القيمة التقديرية للثابت ب من الفصل السابق والتي كانت (٧٥٢ر). بمعنى آخر، اذا زاد الانتاج بمقدار

طن واحـد، فـان سـوف يؤدي الى زيادة تقـديرية بالتكاليف قـدرها (٨٧٩) دينـار بناءاً على خط الانحـدار المتعدد مقارنة مع الزيادة التقـديرية والتي قـدرها (٧٥٢) دينار بناءاً على خط الانحدار الذي تم التعرض إليه في الفصل السابق.

ان السـبب في تبـاين هذين التـقـديرين هـو ان تقـدير تأثيـر كـمـيـة الانتـاج على التكاليف باستـخدام تحليل الانحدار المتعـدد يعمل على بقـاء تأثير سنوات خبرة المدراء ثابت، بينمـا تحليل الانحـدار البسـيط لا يعمل على بقاء تأثير متغير سنوات الخبرة عند المدراء ثابت. وبما ان هذا المتـغير المسـتـقل (عدد سنوات الخبرة) يؤثر على المتـغير التابع (التكـاليف)، فـان هذا يؤدي بالـغالـب الى تحـيز في تقـدير تأثيـر كـمـيـة الانتـاج على التكاليف باستخدام تحليل الانحدار البسيط.

بالطبـع ، يعتبر تحليل الانحدار المتعـدد هذا مناسباً اذا تغيـرت قيـم كـل من $س_{١ر}$ و $س_{٢ر}$ في مدى محدود. فاذا كانت قيـمة $س_{٢ر}$ كبيرة وقيـمة $س_{١ر}$ صغيرة ، فان تحليل الانحدار سـوف يتنبأ بقيـمة سالبـة للتكاليف والذي هو بالتأكيد أمر غير مقبول. ولكن طالما ان تحليل الانحـدار لن يسـتـخـدم في عـمليـة التنبـؤ لقـيم $س_{١ر}$ و $س_{٢ر}$ والتي تقع خـارج مـدى البـيـانات المعطاة في الجـدول (١٣ - ١)، فـانه ليـس هناك أي مـشـكلة. وببساطة، فاننا نفترض، عند استخدام المعادلة (١٣ - ١)، ان تأثير سنوات الخبرة الادارية على متوسط قيمة التكاليف (مع ثبات كمية الانتاج) هو تأثير خطي في مجال محدد فقط.

- فترات الثقة واختبار الفرضيات بالنسبة لـ $ب_١$ و $ب_٢$

Confidence Intervals and Test Hypotheses Concerning B_1 & B_2

وكما هو الحال في تحليل الانحـدار البسيط، فان تحليل الانحدار المتعـدد باستخدام طريقة أقل المربعات له خصائص احصائية مرغوب فيها . وبالتحديد فان القيـم التقديرية أ ، $ب_١$ ، $ب_٢$ هي قيم تقديرية غير متحيزة وقيـم مـستـمرة التـقدير بالنسبة لقيـم خط

الانحدار المتعدد الحقيقية. بالاضافة الى ان نظرية جاوس ماركوف (Gauss Markov Theorem) تقول بأن جميع القيم التقديرية غير المتحيزة للمتغيرات التابعة التي لها اقترانات خطية يكون انحرافها المعياري أقل ما يمكن. وكما هو الحال بالنسبة لتحليل الانحدار البسيط، فان توفر هذه الخصائص المرغوب بها يعتمد على استقلالية المشاهدات وعلى ثبات الانحراف المعياري للتوزيع الاحتمالي المشروط وبغض النظر عن قيم المتغير المستقل. كما انه ليس من الضروري أن يكون التوزيع الاحتمالي المشروط للمتغير التابع توزيعاً طبيعياً. بمعنى آخر، انه ليس من الضروري أن يكون التوزيع الاحتمالي لمقدار الخطأ ($خ_ر$) توزيعاً طبيعياً.

فلو افترضنا ان التوزيع الاحتمالي لمقدار الخطأ ($خ_ر$) موزعاً توزيعاً طبيعياً، فان هذا يعني امكانية حساب فترة الثقة لكل من الثوابت **$ب_١$** أو **$ب_٢$**. وغالباً ما يعتبر هذا غرض مهم من أغراض تحليل الانحدار المتعدد. فعلى سبيل المثال، في حالة المنشأة التي تريد وصف العلاقة ما بين التكلفة وكمية الانتاج وخبرة المدراء، ان الحصول على فترة ثقة للثابت **$ب_١$** يعتبر غرض مهم من أغراض تحليل الانحدار، لأن الثابت **$ب_١$** يمثل معامل المتغير المستقل $س_١$ في خط الانحدار الصحيح لمجتمع الدراسة. ان قيمة الثابت $ب_١$ تمثل القيمة التقديرية لمعامل المتغير المستقل الأول في خط الانحدار الصحيح. وتعتبر القيمة التقديرية لمعامل المستقل الأول مهمة لأنها تقيس تأثير الزيادة في المتغير المستقل $س_١$ بمقدار وحدة واحدة (واحد طن) على القيمة المتوقعة للمتغير التابع ص (التكاليف) في حالة بقاء تأثير خبرة المدراء ($س_٢$) ثابتة. كما يعتبر الحصول على فترة ثقة لمعامل المتغير المستقل $س_٢$ (خبرة المدراء) غرض مهم آخر من أعراض تحليل الانحدار، لأن معامل المتغير المستقل الثاني (**$ب_٢$**) يمثل المعامل الحقيقي للمتغير $س_٢$ في معادلة خط الانحدار للمجتمع. ان قيمة الثابت $ب_٢$ تمثل القيمة التقديرية لمعامل المتغير المستقل الثاني في خط الانحدار الصحيح. وتعتبر القيمة التقديرية لمعامل المتغير المستقل الثاني مهمة لأنها تقيس تأثير الزيادة في المتغير المستقل الثاني بمقدار وحدة واحدة (عدد سنوات الخبرة عند مدراء المنشأة) على القيمة المتوقعة للمتغير التابع ص

(التكلفة)، اذا بقي المتغير المستقل الأول (الكمية) ثابت.

وكما تم الاشارة له في الجزء السابق، فان تحليل الانحدار المتعدد غالباً ما يتم عن طريق أجهزة الحاسوب وليس بالطريقة اليدوية. ونظراً لأهمية تحليل الانحدار المتعدد، فقد تم تطوير برامج معيارية لحساب تحليل الانحدار المتعدد بواسطة طريقة أقل المربعات باستخدام الحاسوب لتقدير قيمة كل من **أ**، **ب$_1$**، **ب$_2$**. وبما ان هذه القيم التقديرية هي قيم احصائية ناتجة عن دراسة العينة، فان هذه القيم تكون خاضعة للأخطاء العينية. فبالاضافة الى حساب القيم التقديرية لكل من **أ** و **ب$_1$** و **ب$_2$** والتي هي قيم كل من أ و ب$_1$ و ب$_2$، فان هذه البرامج المعيارية تقوم بتقدير الانحراف المعياري للثابت ب$_1$ (والذي غالباً ما يطلق عليه بالخطأ المعياري لـ ب$_1$)، وكذلك تقدير الانحراف المعياري للثابت ب$_2$ (والذي غالباً ما يطلق عليه بالخطأ المعياري لـ ب$_2$).

فاذا تم الحصول على نتائج تحليل الانحدار المتعدد بواسطة جهاز الحاسوب، فانه من السهل نسبياً بناء فترة الثقة لكل من **ب$_1$** و **ب$_2$**. وكما تم الاشارة له أعلاه، فانه يمكن الحصول على الخطأ المعياري لكل من ب$_1$ و ب$_2$ في ورقة تحليل النتائج عن طريق الحاسوب. وعليه، فان فترة الثقة لـ **ب$_1$** عند مستوى الثقة ($\propto$) هي:

$$ب_1 \pm t\left(\frac{\propto}{2}\right) S\,(ب_1) \quad \ldots\ldots (13\text{-}6)$$

حيث ان S (ب$_1$) تمثل الخطأ المعياري لـ ب$_1$، بالاضافة الى ان درجات الحرية للاختبار t = ن − ك − ١ (حيث ان ك تمثل عدد المتغيرات المستقلة التي يتضمنها النموذج).

كما ان فترة الثقة لـ **ب$_2$** عند مستوى الثقة ($\propto$) هي:

$$ب_2 \pm t\left(\frac{\propto}{2}\right) S\,(ب_2) \quad \ldots\ldots (13\text{-}7)$$

حيث ان S(ب$_2$) تمثل الخطأ المعياري لـ ب$_2$.

كما أن بعض نتائج تحليل الانحدار المتعدد باستخدام جهاز الحاسوب يقوم بتزويد الخطأ االمعياري للثابت أ. وعليه، فان فترة الثقة للثابت أ عند مستوى الثقة ($\propto$) هي :

$$أ \pm t\,(\frac{\propto}{٢})\,S\,(أ) \quad .\qquad ....\,(١٣-٨)$$

حيث ان S (أ) تمثل الخطأ المعياري لـ أ .

وكذلك اذا تم الحصول على ورقة تحليل جهاز الحاسوب، فانه من السهل اختبار الفرضية الأساسية التي تعتبر قيمة **ب**$_{١}$ أو قيمة **ب**$_{٢}$ مساوية للصفر، في حالة افتراض ان التوزيع الاحتمالي لمقدار الخطأ (خ$_{ر}$) موزعاً توزيعاً طبيعياً. فاذا كانت

$$ب_{١} = صفر$$

فان التوزيع العشوائي للمتغير ب$_{١}$ هو توزيع (t) بدرجات حرية تساوي (ن – ك – ١). فاذا كان اختبار (t) هو اختبار ذو طرفين ومستوى الثقة هو ($\propto$)، فان قانون القرار هو كما يلي :

قانون القرار : أرفض الفرضية الأساسية والتي تعتبر ان قيمة **ب**$_{١}$ تساوي صفر، اذا كانت القيمة المطلقة الاحصائية لـ ب$_{١}$ اكبر من القيمة المعيارية لـ $t\,(\frac{\propto}{٢})$. وقبول الفرضية الأساسية اذا كانت القيمة المطلقة الاحصائية اقل من القيمة المعيارية لـ $t\,(\frac{\propto}{٢})$. علماً بأن درجات الحزية = (ن – ك – ١).

وكذلك نفس الشيء بالنسبة لـ **ب**$_{٢}$. فاذا كانت

$$ب_{٢} = صفر$$

فان قيمة (t) الاحصائية لـ ب$_{٢}$ لها توزيع (t) بدرجات حرية تساوي (ن – ك – ١) . وقانون القرار هو كما يلي:

قانون القرار : أرفض الفرضية الأساسية والتي تعتبر ان قيمة **ب**$_{٢}$ تساوي صفر، اذا كانت القيمة المطلقة الاحصائية لـ ب$_{٢}$ أكبر من القيمة المعيارية لـ $t\,(\frac{\propto}{٢})$. وقبول الفرضية الأساسية اذا كانت القيمة المطلقة الاحصائية اقل من القيمة المعيارية لـ $t\,(\frac{\propto}{٢})$.

علماً بأن درجات الحرية = (ن – ك – ١) .

ان الأمثلة التالية توضح كيفية بناء فترات الثقة لبعض معاملات الانحدار الصحيح او الحقيقي وكيفية اختبار ما اذا كان بعض هذه المعاملات مساوي للصفر عن طريق استخدام ورقة نتائج الحسوب بالنسبة لتحليل الانحدار المتعدد.

مثال (٢) : اذا كان الجزء من حل الانحدار المتعدد باستخدام جهاز الحاسوب للتكلفة بناءاً على كمية الانتاج وسنوات الخبرة عند مدراء المنشأة مزودة بالجدول (١٣ - ٣) (البيانات الأساسية متوفرة في الجدول ١٣ - ١)، وقد تم تسمية كمية الانتاج في هذا الجزء بالمتغير رقم (٢)، بينما تم تسمية سنوات الخبرة عند المدراء بالمتغير رقم (٣). فكما ترى في الجدول (١٣ - ٣)، فان جهاز الحاسوب يطبع قيمة $ب_1$ والخطأ المعياري لـ ($ب_1$) ($S_{(ب_1)}$) وقيمة (t) الاحصائية لـ ($ب_1$) في الصف الأول والمعنون بالرقم (٢) (المتغير الثاني). أما في الصف الثاني والمعنون بالرقم (٣) (المتغير الثالث)، فانه يحتوي على قيمة ($ب_2$) والخطأ المعياري لـ ($ب_2$) ($S_{(ب_2)}$) وقيمة (t) الاحصائية لـ ($ب_2$). وبسبب عملية التقريب في الخطأ، فانه قد وجد هناك بعض الاختلاف ما بين القيم التقديرية لـ أ و $ب_1$ و $ب_2$ المحسوبة بالطريقة اليدوية والمحسوبة بواسطة جهاز الحاسوب. استخدم هذه المعلومات لحساب ما يلي :

١ - فترة الثقة الـ ٩٥٪ لـ **$ب_1$** .

٢ - اختبر ما اذا كانت قيمة **$ب_2$** = صفر ($\propto$ = ٥٪) .

الجدول (١٣ - ٣)

جزء من حل تحليل الارتباط المتعدد باستخدام الحاسب ما بين التكلفة والانتاج والخبرة بواسطة الحاسب.

المتغيرات المستقلة	٢	٣	
المتغيرات التابعة	١		
المتغير	المعامل	الخطأ المعياري	قيمة t
٢	٨٧٩٠ر	٠٣٤٧ر	٢٥ر٣٤٠٠
٣	– ١٧٥٩ر	٠٣٥٣ر	– ٤ر٩٨٠٠
التقاطع	١ر٢٦٤٧٢		

الحل : ١) بما أن الخطأ المعياري لـ $ب_1$ = ٠٣٤٧ر٠ ، فان فترة الثقة الـ ٩٥٪ لـ **$ب_1$** هي :

$$ب_1 \quad \pm\ t\left(\frac{\infty}{٢}\right) S_{(ب_1)}$$

$$٨٧٩ر \pm t\left(\frac{\infty}{٢}\right)(٠٣٤٧ر)$$

بما أن عدد العينة (٩)، وعدد المتغيرات المستقلة هو (٢)، فان

درجات الحرية = ٩ – ٢ – ١

= ٦

وعليه فان t (٠٢٥ر ، ٦) = ٢ر٤٤٧ .

ولهذا، فان فترة الثقة لـ **$ب_1$** هي

$$٨٧٩ر \pm (٢ر٤٤٧)(٠٣٤٧ر)$$

$$٨٧٩ر \pm ٠٨٤٩ر$$

بمعنى آخر، ان فترة الثقة الـ ٩٥٪ للزيادة المتوقعة في التكاليف اذا ازدادت كمية الانتاج بمقدار وطن واحد هي من ٧٩٤ دينار الى ٩٦٤ دينار.

٢) أ) $ف_{.}$: $ب_{٢}$ = صفر

$ف_{١}$: $ب_{٢}$ ≠ صفر

ب) الاختبار هو (t)

جـ) مستوى الثقة ($\propto$) = ٥٪

د) القيمة المعيارية = ٢٫٤٤٧

هـ) القيمة الاحصائية لـ $ب_{٢}$ = ٤٫٩٨٠٠ (من الجدول ١٣ - ٣)

و) القرار : رفض الفرضية الأساسية والتي تعتبر ان قيمة $ب_{٢}$ مساوية للصفر، لأن القيمة الاحصائية المطلقة أكبر من القيمة المعيارية لاختبار (t).

مثال (٣) : اذا كان الجدول (١٣ - ٤) يشير الى نتائج تحليل الانحدار المتعدد باستخدام جهاز الحاسوب، علماً بأن المتغير التابع هو نسبة الزيادة بالانتاجية لشركة ما بين السنوات ١٩٧٢ - ١٩٩٠ . ويوجد هناك ثلاثة متغيرات مستقلة هي : عدد العمال من الشركة الأعضاء مع النقابة في سنة ١٩٧٧ (المتغير الثاني) نسبة الزيادة بالانفاق المضافة على عمل الأبحاث الأساسية في سنة ١٩٧٧ (المتغير الثالث)، ونسبة الزيادة بالانفاق المضافة على عمل الأبحاث التطبيقية والتطوير في سنة ١٩٧٧ (المتغير الرابع). وقد تم حساب خط الانحدار المتعدد بناءاً على (٢٠) مؤسسة صناعية. أوجد ما يلي :

١ - فترة الثقة الـ ٩٥٪ لمعامل الانحدار الحقيقي لنسبة الزيادة بالانفاق على عمل الأبحاث الأساسية

٢ - اختبر ما اذا كان معامل الانحدار الحقيقي لنسبة الزيادة بالانفاق على عمل الأبحاث الحقيقية والتطوير يساوي صفر. (افرض انه مستوى الثقة ($\propto$) = ٥٪) .

الجدول (١٣ - ٤)

جزء من نتائج تحليل الانحدار المتعدد لنسبة الزيادة بالانتاجية للمؤسسة (متغير ١) بناء على نسبة عمالها المسجلين بالنقابة (متغير ٢)، ونسبة الزيادة بالانفاق على الأبحاث الأساسية (متغير ٣)، ونسبة الزيادة بالانفاق على الأبحاث التطبيقية والتطوير (متغير ٤) .

المتغيرات المستقلة	٢	٣	٤
المتغير التابع	١		

المتغير	المعامل	الخطأ المعياري	قيمة (t)
٢	٠٫٠٥٥٩ -	٠٫٠١٤٦	٣٫٨٤٠٠ -
٣	١٫٣٦٢٥	٠٫٤٦٧٠	٢٫٩٢٠٠
٤	٠٫٠٧٢٧	٠٫٠٢٧٦	٢٫٦٤٠٠
التقاطع	٤٫٧٨٨٢٦		

الحل : (١) بما أن قيمة معامل الانحدار التقديرية لنسبة الزيادة بالانفاق على الأبحاث الأساسية هي (١٫٣٦٢٥)، والخطأ المعياري في قيمة معامل الانحدار التقديري هذا هو (٠٫٤٦٧٠)، فان فترة الثقة الـ ٩٥٪ هي :

$$ب_٢ \pm t\,(\frac{\alpha}{٢})\,S_{(ب_٢)}$$

بما أن عدد عناصر العينة هو (٢٠) مؤسسة، وعدد المتغيرات المستقلة هو (٣)، فان درجات الحرية = ٢٠ − ٣ − ١ = ١٦ . وعليه فان قيمة، t (٠٫٠٢٥ ، ١٦) من الجدول المخصص هي (٢٫١٢). وبالتعويض في علاقة فترة الثقة نحصل على ما يلي :

١٫٣٦٢٥ ± (٢٫١٢) (٠٫٤٦٧٠)

١٫٣٦٢٥ ± ٠٫٩٩٠

بمعنى آخر، ان فترة الثقة الـ ٩٥٪ لنسبة الزيادة بالانفاق على الأبحاث الأساسية هي من (٣٧٢٥ر) الى (٣٥٢٥ر٢). ولتفسير هذه النتيجة، افرض ان نسبة الزيادة بالانفاق على عمل الأبحاث الأساسية هي ١٪، فان هذه الزيادة، بالاعتماد على تحليل الانحدار، سوف يؤدي الى زيادة في نسبة الانتاجية ما بين ٣٧٢٥ر الى ٣٥٢٥ر٢، نسبة الى فترة الثقة الـ ٩٥٪ .

٢) أ) $ف_{.}$: $ب_{٢}$ = صفر

$ف_{١}$: $ب_{٢}$ $\neq$ صفر

ب) الاختبار هو (t)

جـ) مستوى الثقة ($\propto$) = ٥٪

د) قيمة (t) المعيارية = ١٢ر٢ (من الجداول المخفضة)

هـ) قيمة (t) الاحصائية = ٦٤ر٢ (من الجدول ١٣ - ٤)

و) القرار : رفض الفرضية الأساسية والتي تعتبر ان قيمة $ب_{٢}$ = صفر، لأن القيمة الاحصائية لاختبار (t) أكبر من القيمة المعيارية لاختبار (t) .

- معامل التحديد المتعدد Multiple Coefficient of Determination

لقد وضحنا في الفصل السابق كيفية استخدام معامل التحديد في قياس درجة تمثيل خط الانحدار البسيط للبيانات. وعند حساب تحليل الانحدار المتعدد، فان معامل التحديد المتعدد هو الذي يستخدم لقياس درجة تمثيل خط الانحدار المتعدد للبيانات وليس معامل التحديد البسيط. ويمكن تعريف معامل التحديد المتعدد كما يلي :

$$ر^{٢} = ١ - \frac{\sum (ص_{ر} - \hat{ص}_{ر})^{٢}}{\sum (ص_{ر} - \bar{ص})^{٢}} \quad ...... (١٣ - ٩)$$

حيث ان $\hat{ص}_ر$ هي قيمة المتغير التابع المتوقعة باستخدام معادلة الانحدار. لهذا، وكما هو الحال في حالة معامل التحديد البسيط الذي تم شرحه في الفصل السابق، فان معامل التحديد هو :

$$ر^2 = \frac{\text{التغير الذي تم تفسيره بواسطة تحليل الانحدار}}{\text{التغير الكلي}} \quad \ldots\ldots (١٣ - ١٠)$$

وهذا يعني ان معامل التحديد ($ر^2$) يقيس ذلك الجزء من التغير الكلي للمتغيرالتابع والذي تم تفسيره بواسطة معادلة الانحدار. وتدعى القيمة الموجبة من الجذر التربيعي لمعامل التحديد المتعدد بمعامل الارتباط المتعدد والذي يرمز له بالرمز ($ر$). كما يستخدم معامل الارتباط في بعض الأحيان لقياس درجة تمثيل خط الانحدار المتعدد للبيانات.

فاذا كان هناك متغيرين مستقلين فقط في تحليل الانحدار المتعدد، فان هناك طريقة سهلة نسبياً لقياس التحديد المتعدد وكما يلي :

$$ر^2 = \frac{ب_1 \sum(س_{1ر} - \bar{س}_1)(ص_ر - \bar{ص}) + ب_2 \sum(س_{2ر} - \bar{س}_2)(ص_ر - \bar{ص})}{\sum ص_ر^2 - \frac{(\sum ص_ر)^2}{ن}} \quad \ldots (١٣ - ١١)$$

فاذا كان عدد المتغيرات المستقلة أكثر من متغيرين، فان تحليل الانحدار غالباً ما يتم عن طريق أجهزة الحاسوب والذي يكون مبرمجاً لحساب معامل التحديد المتعدد أو لمعامل الارتباط المتعدد.

مثال (٤) : استخدم البيانات في الجدول (١٣ - ١) لحساب معامل التحديد المتعدد ما بين تكاليف المنشأة وكمية الانتاج من جهة، وعدد سنوات الخبرة عند المدراء من جهة أخرى. فسّر النتيجة التي تحصل عليها.

الحل : اننا نعلم من حل المثال (١) بان :

قيمة أ = ٠ر٨٨ ،

ب = – ١٨ر ،

Σ(س١ر – سَ١) (صر – صَ) = ٤٦ر٧٨ ،

Σ(س١ر – سَ٢) (صر – صَ) = ٢٩ ،

$$\Sigma \text{ص}_{\text{ر}}^{2} - \frac{(\Sigma \text{ص}_{\text{ر}})^{2}}{\text{ن}} = \text{٣٠٣} - \frac{(\text{٤٩})^{2}}{\text{٩}}$$

= ٣٠٣ – ٢٦٦ر٧٨

= ٣٦ر٢٢

وبتعويض هذه القيم في معادلة معامل التحديد المتعدد، نجد ان قيمة ر² هي كما يلي :

$$\text{ر}^{2} = \frac{\text{٠ر٨٨ (٤٦ر٧٨) + (– ١٨ر) (٢٩)}}{\text{٣٦ر٢٢}}$$

$$\text{ر}^{2} = \frac{\text{٤١ر١٦٦٤ – ٥ر٢٢}}{\text{٣٦ر٢٢}}$$

$$\text{ر}^{2} = \frac{\text{٣٥ر٩٤٦٤}}{\text{٣٦ر٢٢}}$$

$$\text{ر}^{2} = \text{٠ر٩٩}$$

وهذا يعني ان ٩٩٪ من التغير في تكاليف المنشأة الشهرية خلال الفترة التي تغطيها البيانات يمكن تفسيرها بواسطة معادلة تحليل الانحدار المتعدد التي تم الحصول

عليها في المثال رقم (١) .

مثال (٥) : يظهر في الجدول (١٣ ـ ٥) جزء آخر من تحليل الانحدار المتعدد باستخدام جهاز الحاسوب لنسبة الزيادة بانتاجية المؤسسة بناءاً على المتغيرات المستقلة الثلاثة الذي تم وصفهم في المثال رقم (٣) . فسّر القيمة التي تقابل معامل التحديد المتعدد والتي تظهر في الجدول (١٣ ـ ٥) .

جدول (١٣ ـ ٥)

جزء آخر من تحليل الانحدار المتعدد باستخدام جهاز الحاسوب والذي يهتم بنسبة الزيادة بانتاجية المؤسسة.

الخطأ المعياري للتقدير	٦٢٥٥٥ر٠
قيمة (F)	٦٦ر٨
معامل التحديد المتعدد ($ر^٢$)	٦١٩ر٠

الحل : ان القيمة (٦١٩ر) تمثل قيمة معامل التحديد المتعدد . كما تعني هذه القيمة ان معادلة الانحدار المتعدد تستطيع تفسير ما نسبته (٩ر٦١٪) من التغير في نسبة الزيادة بانتاجية المؤسسة .

- تحليل التباين Analysis of Varience

يستخدم تحليل التباين لاختبار الدلالة الاحصائية الكلية (Overall Statistical Significant) لمعادلة خط الانحدار. بمعنى آخر، يستخدم تحليل التباين لاختبار ما اذا كانت جميع معاملات الانحدار في معادلة خط الانحدار تساوي صفراً. ففي مثال المنشأة، فانه من الممكن لادارة المنشأة ان تقوم باختبار ما اذا كان كل من $ب_1$ و $ب_2$ مساوياً للصفر حتى تتمكن من معرفة ما اذا كان هناك علاقة ما بين المتغير التابع وجميع المتغيرات المستقلة معاً. ان تحليل الانحدار يمكن أن يستخدم بهذه الطريقة في تحليل الانحدار البسيط والمتعدد. ففي تحليل الانحدار البسيط، فان النتيجة سوف تكون بالضبط كما هي في اختبار معامل الانحدار ب والذي تم شرحه في جزء سابق.

ان تحليل التباين يتطلب معرفة مجموع التغير في ص والذي يمكن تعريفه على انه مجموع مربع انحرافات القيم ص عن وسطهما الحسابي ($\bar{ص}$). لذلك، نجد ان مجموع التغير في ص يساوي

$$\sum_{ر=1}^{ن} (ص_ر - \bar{ص})^2 \quad .$$

وبالاعتماد على الفصل السابق، نجد ان :

$$\sum_{ر=1}^{ن} (ص_ر - \bar{ص})^2 = \sum_{ر=1}^{ن} (\hat{ص}_ر - \bar{ص})^2 + \sum_{ر=1}^{ن} (ص_ر - \hat{ص}_ر)^2$$

ان المقدار الأول من الجهة اليسرى يمثل التغير المفسّر عن طريق الانحدار، والمقدار الثاني من الجهة اليسرى يمثل التغير غير المفسر عن طريق الانحدار. وعليه، فان

مجموع التغير = التغير المُفسّر + التغير غير المفسر.

وللقيام بتحليل التباين، فإنه لا بد من بناء جدول للشكل العام وكما في الجدول (١٣ - ٦). ان العمود الأول في هذا الجدول يبين نوع أو مصدر التغير (Source of Variation)، والعمود الثاني يبين مجموع المربعات ذات العلاقة (Sum of Squares)،

أما العمود الثالث فانه يبين عدد درجات الحرية (Degrees of Freedom) ذات العلاقة بكل نوع من أنواع التغير. ان درجات الحرية بالنسبة للتغير المفسر (Explained Variation) فانه يساوي عدد المتغيرات المستقلة (ك). اما بالنسبة لدرجات الحرية بالنسبة للتغير غير المفسّر (Unexplained Variation) فانها تساوي (ن – ك – ١). أما بالنسبة للعمود الرابع فانه يبين مربع الوسط (Mean Square) والذي يساوي مجموع المربعات مقسوماً على عدد درجات الحرية. اما العمود الخامس فانه يبين النسبة ما بين مربع الوسط للتغير المفسّر ومربع الوسط للتغير غير المفسر.

جدول (١٣ - ٦)

تحليل التباين للانحدار

مصدر التغير	مجموع المربعات	درجات الحرية	مربع الوسط	نسبة (F)
مفسر بواسطة الانحدار	$\sum(\hat{ص}_{ر} - \bar{ص})^{٢}$	ك	$\frac{\sum(\hat{ص}_{ر} - \bar{ص})^{٢}}{ك}$	$\sum(\hat{ص}_{ر} - \bar{ص})^{٢} / ك$
غير مفسر بواسطة الانحدار	$\sum(ص_{ر} - \hat{ص}_{ر})^{٢}$	ن – ك – ١	$\frac{\sum(ص_{ر} - \hat{ص}_{ر})^{٢}}{(ن - ك - ١)}$	$\sum(ص_{ر} - \hat{ص}_{ر})^{٢}/(ن - ك - ١)$
المجموع	$\sum(ص_{ر} - \bar{ص})^{٢}$	ن – ١		

فاذا كانت الفرضية الأساسية صحيحة (اذا كانت جميع قيم معاملات الانحدار مساوية للصفر)، فان توزيع النسبة سوف يكون توزيع (F) وبدرجات حرية تساوي

(ن – ك – ١). ان قانون القرار لاجراء تحليل تباين مناسب هو :

قانون القرار : (Decision Rule) : ارفض الفرضية الأساسية والتي تعتبر قيم معاملات الانحدار الحقيقية مساوية للصفر اذا كانت نسبة مربع الوسط المفسر الى مربع الوسط غير المفسر تزيد عن قيمة F(∝)، حيث ان ∝ هي مستوى الدلالة المرغوب. وعدا ذلك اقبل الفرضية الأساسية.

وبما ان تحليل الانحدار غالباً ما يتم بواسطة اجهزة الحاسوب، فانه من النادر جداً أن يكون حساب الأرقام في الجدول (١٣ - ٦) بالطريقة اليدوية أمر ضروري. ان مخرجات تحليل الحاسوب غالباً ما تعطي جدول تحليل التباين. فبعض مخرجات الحاسوب تشير لمربع الوسط المفسر بمربع وسط الانحدار (Regression Mean Square)، بينما تشير الى مربع الوسط غير المفسر بمربع الوسط الخطا (Errors Mean Square) أو مربع الوسط المتبقي (Residual Mean Square). ان النسبة ما بين مربع الوسط المفسر الى غير المفسر غالباً ما تمثل قيمة (F) أو نسبة (F). إن الطريقة التي يعرض بها جدول تحليل التباين في مخرجات الحاسوب قد تختلف عن طريقة الجدول (١٣ - ٥)، ولكن المعلومات التي يتضمنها الجدول هي نفسها. وسوف نقوم بتفحص مخرجات أحد البرامج الحاسوبية بالتفصيل فيما بعد من هذا الفصل.

والمثال التالي يوضح تطبيقاً لتحليل التباين.

مثال (٦) : يوضح الجدول (١٣ - ٥) نسبة (F) لتحليل الانحدار المتعدد لنسبة الزيادة بالانتاجية لمؤسسة بناءاً على ثلاثة متغيرات مستقلة تم وصفها في المثال رقم (٣). استخدم هذه النسبة لاختبار الفرضية الأساسية التي تعتبر ان جميع قيم معاملات الانحدار الحقيقية مساوية للصفر (افرض ان مستوى الدلالة (∝) = ٥٪).

الحل : ١) $ف_٠$: $ب_١ = ب_٢ = ب_٣ =$ صفر

$ف_١$: $ب_١ = ب_٢ = ب_٣ \neq$ صفر

٢) $\propto = ٥\%$

٣) الاختبار هو (F)

٤) قيمة (F) الاحصائية = ٨٫٦٦ (من الجدول ١٣ - ٥) .

٥) قيمة (F) المعيارية .

لايجاد قيمة (F) المعيارية، فانه لا بد من تحديد درجات الحرية . ان درجات الحرية هي (٣ و ١٦)، لماذا؟ (لأن عدد المتغيرات (ك) = ٣ ، و (ن – ك – ١) = ١٦) . وعليه، فان

قيمة F (٠٥ر) المعيارية = ٣٫٢٤ (من الجداول المخصصة) .

٦) القرار : رفض الفرضية الأساسية والتي تعتبر ان قيم معاملات الانحدار المتعدد مساوية للصفر.

- مبادىء المتغير الوهمي Dummy-Variable Techniques

يمكن استخدام تحليل الانحدار المتعدد لتحليل تأثير المتغيرات النوعية (Qualitative Varibles) والتي هي تلك المتغيرات التي لا تفترض قيم رقمية، والمتغيرات الكمية (Quantitative Variables) على حد سواء . فعلى سبيل المثال : افرض ان أحد الاداريين يريد تقدير تأثير نسبة الادخار لعائلة بناءاً على متغيرين هما : (١) الدخل السنوي للعائلة، و (٢) هل تمتلك العائلة المنزل التي تسكن به أم تستأجره . ان المتغير المستقل الثاني هو متغير نوعي وليس متغير كمي . ان تحليل الانحدار يمكن أن يتضمن هذا النوع من المتغيرات كمتغير وهمي أو ابكم والذي يعرّف كما يلي :

المتغير الوهمي (الأبكم) : المتغير الوهمي هو المتغير الذي يساوي إما صفر أو واحد.

ففي هذه الحالة، يستطيع الشخص الاداري بناء متغير وهمي ($و_ر$)، والذي يساوي (١) اذا كانت العائلة تمتلك البيت، و (صفر) اذا كانت العائلة لا تمتلك البيت. لهذا، اذا كانت العلاقة ما بين المتغير التابع والمتغير المستقل هي علاقة خطية، فاننا نستطيع افتراض ما يلي :

$$ت_ر = أ + ب_١ د_ر + ب_٢ و_ر + ح_ر \quad ..... (١٣-١٢)$$

حيث ان : $ت_ر$ تمثل قيمة الادخار السنوي للعائلة (ر) .

$د_ر$ تمثل قيمة الدخل السنوي.

ويمكن استخدام طريقة أقل المربعات لتقدير كل من أ و $ب_١$ و $ب_٢$.

ولفهم المقصود من المعادلة (١٣ - ١٢) بشكل واضح وماذا تعني قيمة المتغيرات الوهمية المستخدمة، افرض ان الاداري قد حصل على بيانات من (٢٠) عائلة (بحيث كان عدد العائلات التي تملك بيوتاً لهم يساوي (٦) وعدد العائلات المستاجرة هو ١٤)، كما ان الجدول (١٣ - ٦) يبين هذه البيانات. فاذا تم رسم لوحة الانتشار لقيمة الادخار للعائلة بناءاً على الدخل السنوي، فاننا سوف نحصل على النتائج التي تظهر في الشكل (١٣ - ١). من الشكل (١٣ - ١)، يبدو وبشكل واضح ان العلاقة ما بين الادخار السنوي والدخل السنوي تختلف بالنسبة للعائلات التي تمتك منازلها عن العائلات المستأجرة، وبالتحديد، ان قيمة الادخار للعائلات التي تمتلك منازلها أعلى من ادخار العائلات المستأجرة لكل مستوى من مستويات الدخل. بمعنى آخر، يبدو ان هناك خطين للانحدار، أحدهما للعائلات التي تمتلك منازلها والآخر للعائلات المستأجرة لمنازلها. ان ميل هذين الخطين متساوي ولكن نقطة التقاطع لكلا الخطين مختلفة. افرض

ان هناك خطين للانحدار (بنفس الميل ونقاط تقاطع مختلفة) يظهران في مجتمع الدراسة، كما هو الحال بالنسبة للعينة. ان المعادلة رقم (١٣ - ١٢) سوف تقوم بدمج الانحدار الى خط واحد.

ولمشاهدة ما تقوم به المعادلة (١٣ - ١٢)، افرض ان النموذج الذي يمثل قيمة الادخار للعائلات المستأجرة هو:

$$\text{ت}_{\text{ر}} = \text{أ} + \text{ب}_{١}\,\text{د}_{١} + \text{خ}_{\text{ر}} \text{ ،}$$

وافرض ان النموذج الذي يمثل قيمة الادخار للعائلات المالكة لمنازلها هو :

$$\text{ت}_{\text{ر}} = (\text{أ} + \text{ب}_{٢}) + \text{ب}_{١}\,\text{د}_{١} + \text{خ}_{\text{ر}}$$

حيث ان :

$\text{ب}_{٢}$ تمثل الزيادة بمعدل الادخار بالنسبة للعائلات التي تملك منازلها. وهذا يعني ان المعادلة (١٣ - ١٢) صالحة لكلا النوعين من العائلات، المالكة للمنازل والمستأجرة. لماذا؟ لأن قيمة ($\text{ب}_{٢}$ $\text{و}_{\text{ر}}$) في المعادلة (١٣ - ١٢) تساوي صفر بالنسبة للعائلات المستأجرة، بينما تساوي (ب٢) بالنسبة للعائلات المالكة لمنازلها.

فاذا تم قبول المعادلة (١٣ - ١٢) كنموذج مناسب، فانه يمكن استخدام طريقة اقل المربعات لتقدير كل من أ و $\text{ب}_{١}$ و $\text{ب}_{٢}$ ، وكما يظهر في الجدول (١٣ - ٧). ان خط الانحدار الناتج، والذي يظهر في الشكل (١٣ - ١)، هو

$$\hat{\text{ت}}_{\text{ر}} = -\text{٣٢٠٧ر} + \text{٠٦٧٥ر}\,\text{د}_{\text{ر}} + \text{٨٢٧ر}\,\text{و}_{\text{ر}} \text{ ،}$$

حيث ان :

$\hat{\text{ت}}_{\text{ر}}$ تشير الى قيمة التوفير المتوقعة بآلاف الدنانير للعائلة (ر) باستخدام معادلة

الانحـدار. وبناءاً على هذه المعـادلة، اذا زاد الدخل بمقـدار (١٠٠٠) دينار، فـان ذلك سوف يؤدي الى زيادة بالتوفير قدرها (٦٧٫٥) ديناراً. فاذا تم افتراض ثبات الدخل، فان قيمة التوفير المتوقعة للعائلات التي تملك منزلاً يكون أكثر من توفير العائلات المستأجرة بمقدار ٨٢٧ دينار.

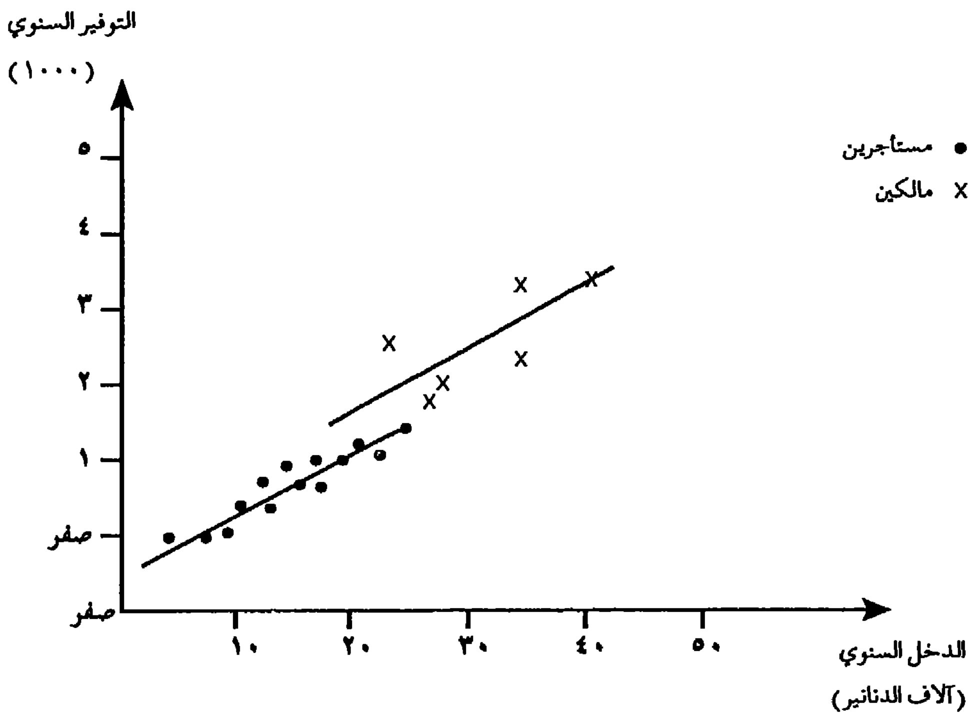
الجدول (١٣ - ٦)

الدخل والتوفير السنوي لـ (٦) عائلات مالكه لمنزلها و (١٤) عائلة مستأجرة

الاسم	التوفير السنوي (١٠٠٠)	الدخل السنوي (١٠٠٠)	مالك / مستأجر	قيمة و
أحمد	١٫٠	٢٠	مستأجر	٠
محمد	١٫٣	٢٤	مستأجر	٠
محمود	٫٧	١٢	مستأجر	٠
علي	٫٨	١٦	مستأجر	٠
خالد	٫٥	١١	مستأجر	٠
عبد الله	٢٫٤	٣٢	مالك	١
قاسم	٫٣	١٠	مستأجر	٠
ابراهيم	٣٫٢	٤٠	مالك	١
حيدر	٢٫٨	٣٢	مالك	١
وائل	٫٠	٧	مستأجر	٠
سائد	٫٣	٩	مستأجر	٠
حامد	٫٠٠	٦	مستأجر	٠
سليم	١٫٠	١٨	مستأجر	٠
سالم	٢٫٠	٢٠	مالك	١
سلامة	٫٤	١٢	مستأجر	٠
عوض	٫٧	١٤	مستأجر	٠
فياض	١٫٥	١٥	مالك	١
نبيه	١٫٦	١٦	مالك	١
سليمان	٫٦	١٥	مستأجر	٠
وليد	٫٦	١٤	مستأجر	٠

الشكل (١٣ ـ ١)

العلاقة ما بين الدخل والتوفير والمستأجرين والمالكين



جدول (١٣ - ٧)

حساب القيم التقديرية لمعالم المجتمع أ و ب١ و ب٢ باستخدام طريقة أقل المربعات وبناءاً على البيانات الواردة في الجدول (١٣ - ٦) .

∑(و ر - وَ)^{٢} = ∑ و^{٢}ر - (∑ و ر)^{٢} / ن = ٦ - ١ر٨ = ٤ر٢

∑(د ر - دَ)^{٢} = ∑ د^{٢}ر - (∑ د ر)^{٢}/ن = ٧٣٧٧ - ٥٨٨٢ر٤ = ١٤٩٤ر٦

∑(ت ر - تَ)^{٢} = ∑ ت ر - (∑ ت ر)^{٢}/ن = ٣٩ر٢٧ - ٢٣ر٥٤٤ = ١٥ر٧٢٦

∑(د ر - دَ) (ت ر - تَ) = ٥١٦ر١ - ٣٧٢ر١٥ = ١٤٣ر٩٥

∑(و ر - وَ) (د ر - دَ) = ١٣ر٥ - ٦ر٥١ = ٦ر٩٩

∑(د ر - دَ) (و ر - وَ) = ١٥٥ - ١٠٢ر٩ = ٥٢ر١

ب١ = [(٤ر٢)(١٤٣ر٩٥) - (٥٢ر١)(١٤٣ر٩٥)] / [(١٤٩٤ر٦)(٤ر٢) - (٥٢ر١)^{٢}] = (٦٠٤ر٥٩ - ٣٦٤ر١٧٩) / (٦٢٧٧ر٣٢ - ٢٧١٤ر٤١) = ر٠٦٧٥

ب٢ = [(١٤٩٤ر٦)(٦ر٩٩) - (٥٢ر١)(١٤٣ر٩٥)] / [(١٤٩٤ر٦)(٤ر٢) - (٥٢ر١)^{٢}] = (١٠٤٤٧ر٢٥٤ - ٧٤٩٩ر٧٩٥) / (٦٢٧٧ر٣٢ - ٢٧١٤ر٤١) = ر٨٢٧

أ = (٢١ر٧ / ٢٠) - ر٠٦٧٥ (٣٤٣ / ٢٠) - ر٨٢٧ (٦ / ٢٠) = -ر٣٢٠٧

(* علماً بان : ∑ ت ر = ٢١ر٧ ، ∑ د ر = ٣٤٣ ، ∑ و ر = ٦)

انه في حالة حذف المتغير الوهمي (و١) من المعادلة وتم استخدام الدخل فقط لتقدير التوفير (الانحدار البسيط)، فان ذلك سوف يؤدي الى تحيز في تقدير قيمة ب١ . وكما يظهر الشكل (١٣ ـ ١)، فان القيمة التقديرية للتوفير بناءاً على الدخل سوف يؤدي الى زيادة بالقيمة المقدرة لـ ب١ لأن العائلات المستأجرة غالباً ما يكون دخلها أقل من دخل العائلات المالكة لمنازلها وبالتالي يكون مقدار التوفير أقل من العائلات المالكة وحتى ولو تساوى دخل العائلات المستأجرة مع دخل العائلات المالكة. وهذا يوضح الحقيقة التي تم التركيز عليها في بداية هذا الفصل والتي تقول : اذا كان المتغير التابع يتأثر بأكثر من متغير مستقل، فان استخدام الانحدار البسيط لتقدير المتغير التابع بناءاً على متغير مستقل واحد غالباً ما يؤدي الى تحيز في تقدير تأثير ذلك المتغير المستقل على المتغير التابع.

وعند استخدام مبدأ التغير الوهمي، فاننا نفترض عدم تأثر معاملات الانحدار الأخرى في معادلة الانحدار بقيمة المتغير الوهمي. فعلى سبيل المثال، لقد تم افتراض ثبات قيمة ب١ في المعادلة (١٣ ـ ١٢) وبغض النظر عن قيمة المتغير الوهمي سواءاً كانت صفر أو واحد. بمعنى آخر، لقد تم افتراض ثبات ميل العلاقة ما بين الدخل والتوفير بالنسبة للعائلات المستأجرة والعائلات المالكة لمنازلها. ان هذا الافتراض قد يكون صحيح أو قد يكون خاطيء (الشكل ١٣ ـ ١ يبين حالة يكون فيها الافتراض صحيح). فاذا كان هذا الافتراض خاطيء، فانه لا بد من ايجاد معادلة خط انحدار منفصلة للعائلات المستأجرة عن معادلة خط انحدار العائلات المالكة.

ـ الارتباط المتعدد Multicollinearity

انه من الخطأ استخدام أي أداة والتي من بينها تحليل الانحدار بشكل عشوائي دون اختبار. لذلك، فانه من المهم جداً اختبار ما اذا كانت شروط تطبيق تحليل الانحدار صحيحة، وكذلك ادراك المشاكل التي يمكن أن تواجه هذا التحليل. ان إحدى المشاكل المهمة التي يمكن أن تظهر في تحليل الانحدار المتعدد هي مشكلة الارتباط المتعدد،

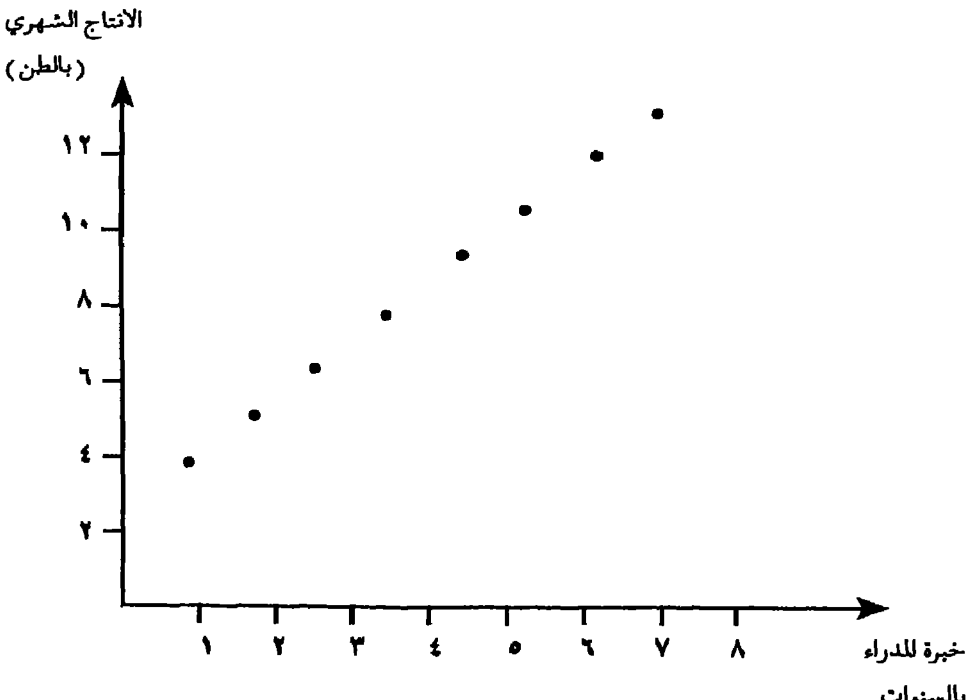
والتي تعرّف كما يلي :

الارتباط المتعدد هو الوضع الذي تكون فيه العلاقة ما بين متغيرين مستقلين أو أكثر علاقة قوية.

ففي مثال المنشأة، افرض ان العلاقة بالماضي ما بين الانتاج وعدد سنوات الخبرة كانت كما هي موضحة بالشكل (١٣ - ٢). فاذا كان الوضع كما هو الحال بالشكل (١٣ - ٢)، فان هذا يعني وجود علاقة ارتباط تامة ما بين متغيرين مستقلين. ففي هذه الحالة، فانه من المستحيل تقدير معاملات الانحدار لكلا المتغيرين المستقلين ($س_١$ و $س_٢$) لأن البيانات لا تزود أية معلومات بالنسبة لتأثير متغير مستقل واحد على فرض بقاء المتغير المستقل الآخر ثابت. وكل ما يمكن ملاحظته هو تأثير كلا المتغيرين المستقلين معاً، على فرض استمرارية وجود علاقة الارتباط ما بين هذين المتغيرين.

الشكل (١٣ - ٢)

العلاقة ما بين المتغيرين المستقلين كمية الانتاج وعدد سنوات الخبرة



ولفهم استحالة تقدير معاملات الانحدار لكلا المتغيرين المستقلين في مثل هذه الحالة، دعنا ننظر الى حالة المنشأة التي تظهر في الشكل (١٣ - ٢). ان تحليل الانحدار يقدر تاثير كل متغير مستقل عن طريق معرفة مدى تاثير هذا المتغير المستقل الواحد على المتغير التابع عند بقاء المتغيرات المستقلة الأخرى ثابتة. لذلك، ففي الحالة التي تظهر في الشكل (١٣ - ٣)، فانه من الصعب جداً تنفيذ هذا التحليل لانه من الصعب جداً لأي شخص من فصل تاثير كمية الانتاج على التكاليف من تاثير سنوات خبرة مدراء المنشأة. فكلما زادت كمية الانتاج، كلما زادت سنوات خبرة مدراء المنشأة. فاذا ما اخذنا بعين الاعتبار طريقة حركة هذين المتغيرين المستقلين، فانه من الصعب جداً ايجاد طريقة لمعرفة مدى تاثير كل متغير مستقل على المتغير التابع بشكل منفصل. ان كل ما نلاحظه في هذه الحالة هو تاثير كلا المتغيرين المستقلين معاً.

فاذا ما كان هناك سبب جيد ومنطقي يجعلنا نعتقد بأن المتغيرات المستقلة سوف تستمر بنفس التاثير على بعضها البعض في المستقبل وكما هو الحال بالماضي، فان الارتباط المتعدد لن يمنعنا (Pervent Us) من استخدام تحليل الانحدار لتقدير المتغير التابع. فبما أن المتغيران المستقلان مرتبطان مع بعضهما البعض ارتباطاً تاماً (Perfectly Correlared)، فان اي منهما يمثل الآخر في التاثير، وبالتالي فاننا نحتاج الى استخدام متغير مستقل واحد في تحليل الانحدار. اما اذا كانت المتغيرات المستقلة لن تستمر بنفس درجة التاثير على بعضها البعض في المستقبل، فان هذا الاجراء سوف يكون خطيراً جداً لأنها لا تأخذ بعين الاعتبار تاثير المتغير المستقل الذي تم استقصاؤه.

بالواقع، فانه من النادر جداً مواجهة حالات تكون فيها المتغيرات المستقلة مرتبطة معاً ارتباطاً تاماً وكما هو الحال في الشكل (١٣ - ٢). ولكن غالباً ما نواجه بحالات تكون فيها درجة ارتباط المتغيرات المستقلة معاً عالية. وعلى الرغم من امكانية تقدير معامل الانحدار لكل متغير مستقل، الا ان هذه التقديرات لا تكون دقيقة. ان الطريقة التي من خلالها نستطيع معرفة وجود الارتباط المتعدد هو عن طريق تقدير معامل الارتباط ما بين المتغيرات المستقلة. فاذا كان معامل الارتباط هذا قريب من +١ أو

(ـ١)، فانه غالباً ما يكون الارتباط المتعدد مشكلة.

ففي بعض الحالات التي يظهر فيها الارتباط المتعدد، فانه من الممكن جداً تغير للتغيرات المستقلة بطريقة يمكن من خلالها التقليل من مشكلة الارتباط المتعدد. أما اذا كان من الصعب تجنّب الارتباط المتعدد بهذه الطريقة، فانه لا يوجد هناك بديل سوى الحصول على بيانات جديدة لا تحتوي على ارتباط قوي ما بين المتغيرات المستقلة.

ـ الانحدار المتعدد وبرامج الحاسوب

Computer Programs and Multiple Regression

ان ظهور (Advent) التقنية الحاسوبية قد أدى الى تقليل الجهد (Effort) والتكلفة (Expense) المطلوبة لحساب الارتباط المتعدد للمتغيرات المستقلة الكثيرة بشكل واضح (Marked Reduction). فقبل ظهور الحاسوب، كان الجهد المطلوب لحساب الانحدار المتعدد لعدد قليل من المتغيرات المستقلة كبيراً جداً. أما الآن فان حساب الانحدار المتعدد سهل نسبياً (Relatively Simple). انه من المهم جداً الآن لك أن تعرف وتعلم المعلومات المختلفة والمطبوعة بواسطة جهاز الحاسوب والأشكال التي تظهر بها. وبما أن البرامج المقننة التي تستخدم لحساب تحليل الانحدار كثيرة، فانه لا يوجد هناك شكل وحيد أو قائمة بالفقرات التي يطبعها جهاز الحاسوب. ولكن جميع الأشكال المختلفة لمخرجات جهاز الحاسوب متماثلة، وعليه فانه من المفيد جداً النظر الى أحد أشكال المخرجات بشيء من التفصيل.

ان الجدول (١٣ ـ ٨) يظهر أحد أشكال مخرجات الحاسوب لتحليل الانحدار لنسبة الزيادة في الانتاجية لأحد الصناعات خلال الفترة ١٩٧٢ ـ ١٩٩٠ بناءاً على ثلاثة متغيرات مستقلة هي :

١ ـ نسبة عمال الصناعات الأعضاء في النقابة (المتغير رقم ٢) .

٢ ـ نسبة الزيادة بالانفاق على الأبحاث الأساسية (المتغير رقم ٣) .

٣ ـ نسبة الزيادة بالانفاق على الأبحاث التطبيقية (متغير رقم ٤) .

اننا سوف نقوم الآن على معرفة ما تعنيه الصفوف الأربعة العليا من هذه المخرجات والتي ظهرت في الجدول (١٣ ـ ٤) . أما بالنسبة للصفوف الستة الأخرى من الجدول (١٣ ـ ٨)، فانها تظهر نتائج تحليل التباين للانحدار المتعدد. ان التغير المفسر بواسطة الانحدار (The Varitation Explained by the Regression) يساوي (١٧١٠٤ر١٠). والخطأ في مجموع المربعات (The Error Sum of Squares) يساوي (٢٦٠٩٦ر٦). ان مربع وسط الانحدار (The Regression Mean Square) يساوي التغير المفسر بالانحدار مقسوماً على درجات الحرية (عدد المتغيرات المستقلة). لهذا، فان مربع وسط الانحدار في هذه الحالة يساوي (١٧١٠٤ر١٠ ÷ ٣ ، أو ٣٩٠٣٥ر٣). ان مربع وسط الخطأ هنا (The Error Mean Square) هو الخطأ في مجموع المربعات مقسوماً على درجات الحرية له (والتي تساوي عدد المشاهدات – عدد المتغيرات المستقلة – ١). لهذا، فان مربع وسط الخطأ في هذه الحالة يساوس (٢٦٠٩٦ر٦ ÷ ١٦ (٢٠ – ٣ – ١) = ٣٩١٣١ر٠). أما قيمة F فانها تظهر في الصف قبل الأخير من الجدول (١٣ ـ ٨) .

أما التعبير الذي لم يفسر من الجدول (١٣ ـ ٨) هو الخطأ المعياري للتقدير (The Standard Error of Estimate). وكما هو الحال بالنسبة للانحدار البسيط، فان الخطأ المعياري للتقدير هو التقدير للانحراف المعياري بالنسبة للتوزيع الاحتمالي للمتغير التابع عندما تكون جميع المتغيرات المستقلة ثابتة. لهذا، فان الخطأ المعياري للتقدير يقيس مقدار انتشار المشاهدات عن خط الانحدار. وفي هذه الحالة، فان مقدار الخطأ المعياري للتقدير يساوي (٦٢٥٥٥ر)، والذي يعني ان الانحراف المعياري للفرق ما بين نسبة الزيادة في انتاجية الصناعة الحقيقية والمتوقعة بواسطة خط الانحدار يساوي (٦٢٥٥٥ر) .

الجدول (١٣ ـ ٨)

نتائج تحليل الانحدار المتعدد لنسبة الزيادة في انتاجية لأحد الصناعات

المتغيرات المستقلة	٢	٣	٤
المتغير التابع	١		

المتغير	المعامل	الخطأ المعياري	قيمة t
٢	ـ ٠٥٥٩ر	٠٤٦ر٠	ـ ٨٤٠٠ر٣
٣	٣٦٢٥ر١	٤٦٧٠ر٠	٩٢٠٠ر٢
٤	٠٧٢٧ر	٠٢٧٦ر٠	٦٤٠٠ر٢

التقاطع	٧٨٨٢٦ر٤
الانحدار	Regression
درجات الحرية	٣
مجموع المربعات	١٧١٠٤ر١٠
مربع الوسط	٣٩٠٣٥ر٣
الخطأ	Error
درجات الحرية	١٦
مجموع المربعات	٢٦٠٩٦ر٦
مربع الوسط	٣٩١٣١ر٠
الخطأ المعياري للتقدير	٦٢٥٥٥ر٠
قيمة F	٦٦ر٨
مربع الارتباط المتعدد	٦١٩ر

أسئلة

س١ : يريد باحث تقدير العلاقة ما بين الانفاق السنوي على الطعام من جهة وعدد افراد الأسرة والدخل للأسرة لعائلات من منطقة لواء المزار الجنوبي. وقد حصل هذا الباحث على بيانات لعشرة عائلات تم اختيارهم عشوائياً.

الانفاق السنوي على الطعام (بالمئات)	عدد أفراد الأسرة	الدخل السنوي للأسرة (بالمئات)
٨٠ر	٤	٢٦
١ر٥٠	٦	٣٠
٤٠ر	٤	١٥
٢ر٤٠	٦	٤٢
٣ر٨٠	٦	٥٣
٢ر٠٠	٧	٣٢
٣ر٥٠	٨	٤٠
٢ر٩٠	٧	٤٠
٣ر٠٠	٨	٣٠
٤ر٠٠	٦	٥٦

أ ـ أوجد خط الانحدار المتعدد للعينة.

ب ـ ما هو التاثير التقديري الناتج عن زيادة عدد افراد الأسرة بفرد واحد على الانفاق السنوي على الطعام.

جـ ـ ما هو التاثير التقديري الناتج عن زيادة الدخل بمقدار (١٠٠) دينار على الانفاق السنوي على الطعام.

د ـ ما هو الانفاق السنوي على الطعام لعائلة تمتلك أربعة أطفال ودخلها السنوي (٣ر٥٠٠) دينار.

س٢ : تريد إحدى الجامعات تقدير العلاقة ما بين معدل علامات الطالب من جهة ومعدل علاماته بالثانوية العامة وعلامة الـ(GMAT) من جهة أخرى. فاذا كانت هذه المتغيرات لـ(١٢) طالب كما يلي :

معدل علامات الجامعة	معدل الثانوية العامة	علامة ألـ GMAT
٢ر٥	٣ر٠٠	٥٨٠
٣ر٠٠	٣ر٢٠	٦٠٠
٣ر٠٠	٣ر٥٠	٦٨٠
٣ر١٠	٣ر٤٠	٦٥٠
٣ر٢٠	٣ر٨٠٠	٧٢٠
٣ر٦٠	٤ر٠٠	٧٥٠
٣ر٥٠	٣ر٧٠	٧٥٠
٣ر٥٠	٣ر٨٠	٧٦٠
٣ر٦٠	٣ر٩٠	٧٥٠
٣ر٩٠	٣ر٨٠	٧٤٠
٣ر٨	٤ر٠٠	٧٨٠
٤ر٠٠	٤ر٠٠	٧٨٠

أ ـ أوجد خط الانحدار المتعدد للعينة.

ب ـ ما هو التأثير التقديري الناتج عن زيادة معدل الثانوية بمقدار (١ر٠٠) على معدل الطالب الجامعي.

ج ـ ما هو التأثير التقديري الناتج عن زيادة (١٠٠) علامة في علامة الـ (GMAT) على معدل الطالب الجامعي.

د ـ ما هو معدل الطالب الجامعي التقديري الذي كان معدله بالثانوية العامة (٣ر٦) وعلامته بالـ (GMAT) هي (٧٤٠).

س٣ : يهتم اقتصادي زراعي في تقدير العلاقة ما بين الأسمدة النيتروجينية والأسمدة الفوسفاتية من جهة وحجم الانتاج للذرة من جهة أخرى. فاذا استخدم عدد أكياس الذرة المنتجة للدونم الواحد كمتغير تابع (متغير رقم ١)، ووزن السماد النيتروجيني بالكيلوغرام للدونم الواحد (متغير رقم ٢) ووزن السماد الفرسفاتي بالكيلرغرام للدونم الواحد (متغير رقم ٣)، وكانت عدد المشاهدات ٢٥. كانت مخرجات جهاز الحاسوب كما يلي :

المتغيرات المستقلة	٢	٣	
المتغير التابع	١		
المتغير	المعامل	الخطأ المعياري	قيمة t
٢	٢٩ر	١١ر	٢ر٦٤
٣	٦٧ر	٢٣ر	٢ر٩١
التاطع	١٩ر		

أوجد ما يلي :

أ ـ احسب فترة الثقة ٩٠٪ لمعامل الانحدار الصحيح لوزن الأسمدة النيتروجينية

بالكيلوغرام للدونم الواحد .

ب ـ احسب فترة الثقة ٩٥٪ لمعامل الانحدار الصحيح لوزن الأسمدة الفوسفاتية بالكيلوغرام للدونم الواحد .

س٤ : بالاعتماد على البيانات الواردة في السؤال رقم (١)، احسب معامل التحديد ـ فسّر اجابتك .

س٥ : بالاعتماد على البيانات الواردة في السؤال رقم (٢)، احسب معامل التحديد . فسّر اجابتك .

س٦ : بالاعتماد على البيانات الواردة في السؤال رقم (١)، استخدم هذه البيانات لاختبار الفرضية الأساسية التي تعتبر ان جميع قيم معاملات الانحدار الحقيقة مساوية للصفر، عند مستوى الدلالة ٥٪ .

س٧ : بالاعتماد على البيانات الواردة في السؤال رقم (٢)، استخدم هذه البيانات لاختبار الفرضية الأساسية التي تعتبر ان جميع قيم معاملات الانحدار الحقيقية مساوية للصفر، عند مستوى الدلالة ١٪ .

س٨ : بالاعتماد على البيانات الواردة في السؤال رقم (١)، هل هناك دليل لوجود الارتباط المتعدد ؟

س٩ : بالاعتماد على البيانات الرواردة في السؤال رقم (٢)، هل هناك دليل لوجود الارتباط المتعدد ؟

س١٠ : أ) استخدم البيانات أدناه لحساب الانحدار المتعدد ما بين عدد الجرائم من جهة ونسبة البطالة والمتغير الوهمي الذي يمثل السنة التي حدثت بها الجريمة.

المدينة	عدد الجرائم	نسبة البطالة (٪)	السنة
أ	١٠٠	٥	١٩٩٠
ب	١٢٠	٧	١٩٩٠
ت	١٤٠	٨	١٩٩٠
ث	١٧٠	٩	١٩٩٠
جـ	١١٠	٥	١٩٩٢
ح	١٦٠	٨	١٩٩٢
خ	٢٠٠	٩	١٩٩٢
د	١٢٠	٦	١٩٩٢
ر	١٣٠	٧	١٩٩٢
ز	١٥٠	٧	١٩٩٢

ب) بالاعتماد على نتائج الفرع (أ)، ما هي القيمة التقديرية للفرق بالوسط ما بين سنة ١٩٩٠ وسنة ١٩٩٢ في عدد الجرائم للمدينة، على فرض ثبات نسبة البطالة. وما هو التأثير المتوقع الناتج عن زيادة نسبة البطالة بنسبة ١٪ ، على فرض ثبات السنة.

الفصل الرابع عشر

كتابة تقرير البحث

Reporting Research Findings

كتابة تقرير البحث

Reporting Research Findings

- مقدمة

ان آخر خطوة من خطوات عملية البحث هي الاعداد والعرض لنتائج البحث في تقرير يسمى بتقرير البحث، حتى يتسنى للجهة المموّلة للدراسة من الاطلاع على نتائج البحث من أجل الاستفادة من نتائج البحث في عملية اتخاذ القرار. وفي بعض الأحيان، قد يكون الهدف من كتابة تقرير البحث هو من أجل ارساله الى مجلة أو دورية لغايات النشر، أو كمتطلب لاستكمال متطلبات الحصول على شهادة أو درجة علمية يسعى الشخص الباحث الحصول عليها. وتختلف متطلبات كتابة تقرير البحث من مؤسسة الى أخرى، أو من جامعة الى أخرى، أو من دورية الى أخرى. وغالباً ما يكون الهدف من تحديد مواصفات للكتابة هو للتأكد من سلامة اللغة ولضمان تحقيق الفائدة من اجراء البحث، والتأكد من صحة المعلومات الواردة في التقرير حتى تتمكن الجهة الممولة من الاستفادة من نتائج الدراسة في عملية اتخاذ القرار. كما تهدف متطلبات كتابة التقرير التأكد من ملائمة التنظيم وكفاءته في ايصال المعلومة والمعرفة بسهولة ويسر للشخص القارىء.

تعتبر هذه الخطوة في بعض الأحيان من أكثر الأجزاء اهمية في عملية البحث. فاذا كانت كتابة تقرير البحث ضعيفة أو ركيكة أو مربكة، فان هذا يعني ان كل الوقت والجهد الذي انقضى في تجميع وتحليل البيانات كان بدون فائدة. فنجد ان بعض الأشخاص يمتلكون الميل الى الكتابة والرغبة في تعلّم وامتلاك مهاراتها ويتمتعون عند القيام بكتابة تقرير البحث. بينما نجد أن البعض الآخر يفضل القيام باجراءات البحث

ويراها أسهل عليه من كتابة تقرير يصف فيه الاجراءات التي قام بها والنتائج التي حصل عليها. ولعلنا نجد في ذلك ما يبرّر ضرورة وجود شخص له اهتمامات بالكتابة ضمن فريق البحث اذا كان البحث يتطلب انجازه وجود فريق عمل.

لهذا، فان الهدف من كتابة هذا الفصل هو لاعطاء بعض الخطوط الاسترشادية والتي يمكن ان تؤخذ بعين الاعتبار عند كتابة تقرير البحث، والذي غالباً ما يكون موجهاً لجهة محددة. كما يجب ان نهتم في هذا المجال بالتقارير الشفهية والكتابية، لأنه غالباً ما يتم اعداد هذين النوعين من التقارير عند الانتهاء من أي مشروع بحث.

ان الجانب الوحيد الذي يدركه معظم المدراء من جوانب عملية البحث هو إما التقرير الشفهي أو الكتابي أو كليهما. لذلك، فان تقييمهم للأسابيع أو الأشهر التي استخدمها الشخص الباحث تعتمد على عملية العرض. وعليه، فانه من الممكن ان تكون أفضل منهجية بحث في العالم عديمة الفائدة بالنسبة للمدراء أو المسؤولين، اذا لم يتمكن الباحثون من عرض ما هو مهم بالبحث من خلال تقرير البحث. لهذا، فاننا سوف نقوم بعرض بعض النصائح بالنسبة لكيفية عرض نتائج البحث بوضوح وبطريقة أو أسلوب ممتع.

- خطوط استرشادية لكتابة التقرير . Written Report Guidelines

يوافق مجموعة كبيرة من الباحثين المهرة عند كتابة تقرير البحث على وجود سلسلة من الخطوط الاسترشادية والتي يجب ان تتبع عند كتابة تقرير البحث. وهذه الخطوط هي :

١ - اعتبار جمهور القرآء . Consider the Audience

يجب على الشخص الباحث اعتبار جمهور القراء عند البدء بكتابة تقرير البحث وذلك عن طريق اعتبار النقاط التالية :

أ ـ كتابة تقرير واضح. Make the report clear

ب ـ استخدام كلمات مستخدمة عند الجمهور. Use only words familiar to the readers

جـ تسهيل عمليات المقارنة من خلال استخدام الرسومات التوضيحية.

Use figures to make the comparison easier.

د ـ استخدام النسب والرسومات التقديرية والرتب للتسهيل والوضوح.

Use percentages, rounded-off figures, or ranks.

هـ ـ وضع البيانات وبدقة في جدول اما في متن البحث او في الملحق.

Put the exact data in a table within the text or in the appendix.

و ـ استخدام الرسومات المساعدة في أي مكان يمكن من شأنه أن يساعد في توضيح البيانات المعروضة.

Use graphic aids (charts, graphs, etc.) whenever they help clarify the presentation of data.

٢ ـ تسمية أو تحديد المعلومات المطلوبة. Address the information needs

يجب على الشخص الباحث ان يتذكر ان تقرير البحث مصمم لايصال المعلومات الى متخذي القرارات (Decision makers). لذلك، لا بد للشخص الباحث ان يتأكد من ربط (Relate) نتائج الباحث (Research Findings) مع اهداف البحث (Objectives of Mannagement) وبوضح (Clearly).

٣ ـ مراعاة الدقة والكمال Be Concise and Complete

يجب على الشخص الباحث أن يراعي ان معظم المدراء (Most Managers) لا

يرغبون دراسة تفاصيل (Details) مشروع البحث (Research Project) لعدم امكانية توفير الوقت في بعض الأحيان، أو لعدم توفر الرغبة في بعض الأحيان الأخرى. ان معرفة ما يجب تضمينه من معلومات لمتن البحث وما يجب حذفه (Leave Out) هي مهمة ليست بالسهلة (Difficult Task). ان عملية تضمين او حذف المعلومات تعتمد على معرفة الشخص الباحث بالمعلومات المطلوبة لمتخذ القرار عند كتابة التقرير.

٤ ـ الموضوعية Be Objective

غالباً ما يواجه الشخص الباحث على الاقل موقفاً واحداً والذي فيه يكون قبول النتائج من قبل جمهور القراء ليس بالامر السهل. فالنتائج قد تتعارض (Conflict) مع خبرة متخذ القرار (Experience of Decision Maker) وحكمه (Judgment)، او يكون تاثير النتائج على حكمة (Wisdom) القرارات السابقة (Previous Decisions) غير محبب أو مرغوب به (Unfavorable). ان ظهور مثل هذه الظروف (Circumstances) قد يكون عامل قوي لتحريف (Slant) تقرير البحث، لجعل نتائج مشروع البحث اكثر قبولاً (More Acceptable) الى الادارة. ان الباحثين المهرة غالباً ما يعرضوا نتائج البحث بطريقة موضوعية (Objective Manner) وبدون تحيز (Without Bias)، ومن ثم الدفاع (Defend) عن نتائج البحث في حالة التحدي (Challenge) من قبل اي قارىء.

٥ ـ أسلوب الكتابة Style

يعتبر أسلوب الكتابة احد المواضيع المتعلقة بمساق طرق الاتصالات، ولكن سوف نقوم هنا بعرض القليل من الأفكار (Few Tips) التي تساعد على كتابة تقرير سهل القراءة. وهذه الافكار هي :

أ ـ استخدام الكلمات والجمل القصيرة. Use short words and sentences

ب ـ الدقة في التعبير. Be Concise

جـ ـ اعتبار مظهر التقرير. Consider Appearance

ان شكل التقرير في بعض الأحيان قد يؤثر على نوعية تقرير البحث. فعلى سبيل المثال، قد تساعد المسافات ما بين الأسطر على تسهيل قراءة التقرير الطويل. وكذلك، فان استخدام الرسومات البيانية قد يوضح بعض الأفكار الاحصائية.

د ـ تجنب الصيغ او الأفكار المبتذلة Aviod Cliches

هـ ـ الكتابة في صيغة الحاضر. Write in the present tense

و ـ وضع بعض الاقتباسات (Quotes) من المستجوبين (Respondents) ضمن تقرير البحث لاضفاء نوع من المتعة على التقرير لجعله أكثر قابلية للقراءة.

ـ أنواع الأبحاث Types of Research

يمكن تقسيم الأبحاث من حيث الوقت اللازم لاكمال البحث الى الأنواع التالية :

١ ـ التقرير Report

ان الهدف من اعداد التقرير بشكل عام هو تلخيص كتاب أو مجموعة من الكتب حول موضوع معين، أو وصف ظاهرة اما علمية أو اجتماعية كان قد أحس بها الشخص الباحث.

٢ ـ الورقة الفصلية Term Paper

تتضمن الورقة الفصلية القيام ببحث قصير من أجل الحصول على معلومات حول موضوع معين. وغالباً ما يكون الهدف من وراء القيام بالورقة الفصلية هو لتقديمه

كجزء رئيسي من متطلبات النجاح لبعض المساقات الدراسية. بمعنى آخر، ان هذا النوع من الأبحاث غالباً ما يقوم به الطلبة من أجل اجتياز بعض المواد التي تتطلب اعداد ورقة فصلية. لهذا، قد تعالج هذه الأوراق البحثية بعض المشاكل التي لا تحتاج الا الى المادة المكتبية، أو قد تعالج هذه الأوراق البحثية بعض المشاكل التي تحتاج الى تجميع بيانات أولية من خلال الدراسات الميدانية. وعليه، فان تكليف الطالب بالقيام بمثل هذا النوع من الأبحاث قد يساعده في استفادة بعض المعلومات المكملة للمادة. كما تعلّم الطالب الاعتماد على نفسه في تجميع المعلومات وتحمل المسؤولية والفهم والانتقاد البناء والدقة في الاقتباس الى غير ذلك من الفوائد.

٣ ـ رسالة الماجستير Dissertation

ان الهدف من القيام بمثل هذا النوع من الأبحاث هو انجاز أحد المتطلبات الرئيسية للحصول على درجة الماجستير. لذلك، غالباً ما يكون هذا النوع من الأبحاث أطول وأعمق وأشمل من الأبحاث السابقة. كما ان انجاز هذا النوع من الأبحاث يتطلب من الشخص الباحث أن يقوم بمعالجة موضوع جديد لم يتم بحثه من قبل أي بحث آخر.

٤ ـ أطروحة الدكتوراة Thesis

ان الهدف من القيام بمثل هذا النوع من الأبحاث وكما هو الحال بالنسبة لرسائل الماجستير هو انجاز أحد المتطلبات الرئيسية للحصول على درجة الدكتوارة. وبما ان درجة الدكتوراة أعلى مستوى من درجة الماجستير، فان هذا يتطلب من الشخص الباحث عند القيام باعداد اطروحة الدكتوراة معالجة مشكلة معينة تحتاج الى دراسة أوسع وأعمق وأشمل من تلك المشكلة التي تعالجها رسالة الماجستير. لهذا، فان انجاز اطروحة الدكتوراة يتطلب الاستعانة بمصادر ومراجع

متعددة، والتعمق بالموضوع، والدقة في تصميم البحث، والبراعة في التحليل والتقييم لمادة موضوع البحث، وكذلك المقدرة على التعبير عن نتائج البحث بامانة وصدق وموضوعية عند البدء باعداد تقرير نتائج البحث. وغالباً ما يتطلب انجاز اطروحة الدكتوراة الى اجراء دراسات ميدانية في العلوم الاجتماعية أو اجراءات تجارب مختبرية في العلوم الطبيعية، من أجل التوصل الى نتائج واستنتاجات واقتراحات وتوصيات ذات قيمة بالنسبة لحل المشكلة قيد البحث وتوسيع حدود المعرفة الانسانية بالنسبة لموضوع مشكلة الاطروحة.

- شكل التقرير. Report Format

على الرغم من عدم وجود شكل قياسي للتقرير بحيث يمكن استخدامه في جميع الحالات، الا انه هناك نوع من الاجماع من معظم الباحثين على قبول الشكل الأساسي التالي لمعظم مشاريع الأبحاث.

١ - صفحة العنوان. Title Page

٢ - صفحة الشكر والتقدير. Acknowledgement

٣ - قائمة المحتويات. Table of Contents

٤ - قائمة الجداول. Table of Tables

٥ - قائمة الأشكال. Table of Figures (Graphs)

٦ - ملخص الدراسة. Abstract

٧ - متن الدراسة. Body

٨ - الخلاصة والتوصيات. Conclusions and Recommedations

٩ - المراجع. References

١٠ - الملحقات. Appendixes

١ ـ صفحة العنوان Title Page

يجب ان تتضمن هذه الصفحة على عنوان يتمكن من نقل ما يلي :

أ ـ جوهر الدراسة. Essenc of the Study

ب ـ تاريخ الدراسة. Date of the Study

جـ ـ اسم الجهة أو المؤسسة المقدمة للبحث. Name of submitting the report

د ـ اسم الجهة المستلمة للبحث Name of recipient

هـ ـ اذا كانت نتائج البحث سرية (Confidential)، فيجب كتابة اسم الشخص الذي يجب أن يستلم البحث على صفحة العنوان.

١ ـ صفحة الشكر والتقدير Acknowledgement

غالباً ما تتضمن هذه الصفحة الشكر والتقدير لكل من ساهم في انجاز البحث قيد الدراسة مثل الجهة الممولة للبحث ورجال المكتبة والمحللون الاحصائيون ورجال الحاسوب وما الى ذلك. ويفضل أن توضع هذه الصفحة بعد صفحة العنوان لكي يتمكن القارىء من معرفة اسم كل الجهات التي ساهمت في اعداد هذا البحث.

وفي بعض الأحيان، قد نجد ان هذه الصفحة قد تكون مسبوقة بصفحة الاهداء والتي غالباً ما تتضمن أسماء الأشخاص أو الجهات التي يرغب الشخص الباحث من اهداء عمله اليهم مثل الوالدين أو الزوجة أو الأبناء. لذلك، فان الفرق ما بين صفحة الاهداء وصفحة الشكر والتقدير هو ان الأشخاص اللذين تُكتب أسماؤهم في صفحة الشكر والتقدير غالباً ما يكونوا قد قدموا خدمة للشخص الباحث، بينما الأشخاص اللذين ترد أسماؤهم في صفحة الاهداء ليس من الضروري أن يكونوا قد قدموا خدمة للشخص الباحث.

٣ ـ قائمة المحتويات Table of Tables

تهدف هذه القائمة الى عرض جميع المواضيع التي تم تغطيتها في متن البحث وبتسلسل مع أرقام الصفحات التي تم عرضها بها. ان الغرض من ذلك العرض هو مساعدة الشخص القارىء في ايجاد الجزء موضع الاهتمام بالنسبة له بسهولة ويسر. كما تتضمن هذه القائمة المراجع والمصادر والملحقات والفهرس ان وجدت.

٤ ـ قائمة الجدوال والأشكال Table of Tables and Figures

تهدف هذه القائمة الى عرض اسم ورقم الصفحة لكل جدول أو شكل موجود في متن البحث، لتسهيل عملية الحصول على أي جدول أو شكل يكون موضوع اهتمام بالنسبة للشخص القارىء بسهولة ويسر دون عبث المرور بجميع صفحات متن البحث. ويمكن وضع هذه القائمة مع قائمة المحتويات أو في صفحة مستقلة. كما لا يوجد هناك أي مانع من وضع قائمة الجدوال في صفحة مستقلة عن قائمة الأشكال.

٥ ـ ملخص الدراسة Abstract

يتطلب كتابة ملخص الدراسة التركيز والدقة عند كتابة الجمل لتوضيح ما هو مهم في تقرير البحث. وغالباً ما يكون عدد صفحات الملخص من صفحتين الى ثلاثة صفحات، كما انها وجودها هو أمر ضروري لمعظم تقارير الأبحاث. وبما أن معظم متخذي القرارات لا يقومون الا بقراءة هذا الجزء من البحث، فانه من المهم جداً كتابة هذا الجزء بدقة وبطريقة جيدة.

ان ملخص الدراسة لا يعتبر صورة مصغرة (Miniature) عن التقرير الأصلي. انما الهدف منه هو تزويد الشخص متخذ القرار بأهم نتائج البحث ذات التاثير القوي (Most impact) على القرار المراد اتخاذه. وبما ان ملخص الدراسة يكتب بالتحديد الى متخذي القرارات، فانه لا بد لهذا الملخص ان يكون قادر على مساعدتهم في اتخاذ

القرارات. لذلك، يجب أن يحتوي ملخص الدراسة على ما يلي :

١ ـ أهداف مشروع البحث. Objectives of the Research Project

٢ ـ طبيعة مشكلة القرار. Nature of the Decision Problem

٣ ـ النتائج المهمة. Key Results

٤ ـ الخلاصة. Conclusions

٥ ـ التوصيات. Recommendations for Actions

٦ ـ متن الدراسة Body of the Study

يتضمن متن الدراسة تفصيلات مشروع البحث. وغالباً ما يتضمن متن الدراسة المواضيع التالية :

أ ـ المقدمة. Introduction

ب ـ أدبيات الدراسة. Literature Review

جـ ـ منهجية البحث. Research Methodology

د ـ النتائج. Results

هـ ـ محددات البحث. Limitations

أ ـ المقدمة Introduction

ان الهدف من كتابة المقدمة هو لتزويد الشخص القارىء بخلفية عن المعلومات المطلوبة (Information needed) لفهم ما يتبقى (The Remainder) من تقرير البحث. ان طبيعة المقدمة تكون مشروطة بنوعية جمهور القراء ومدى معرفتهم بمشروع البحث. فكلما تنوع جهور القراء، كلما زادت شمولية المقدمة. يجب على المقدمة أن تشرح وبوضوح طبيعة مشكلة البحث وأهداف الدراسة.

لذلك، يجب على المقدمة ان تقوم بتزويد معلومات كافية عن الظروف المحيطة بمشكلة القرار. كما يجب التعرض الى اي بحث سابق كان قد تم عن مشكلة البحث ومراجعته.

ب ـ أدبيات الدراسة Literature Review

ان الهدف من مراجعة أدبيات الدراسة بالنسبة للمشكلة قيد البحث هو ما يلي :

١ ـ تعميق فهم الشخص الباحث حول المشكلة قيد الدراسة.

٢ ـ المساعدة في تشخيص المشكلة قيد البحث ووضع أهدافها.

٣ ـ العمل على تزويد معلومات للمقارنة.

٤ ـ المساعدة في ايجاد الطريقة الأمثل لتحليل المشكلة قيد الدراسة.

٥ ـ التعرف على ما تمّ التوصل اليه من معرفة حول الموضوع الذي تقع ضمنه المشكلة قيد البحث.

٦ ـ للتأكد من عدم بحث المشكلة قيد البحث بالسابق.

جـ ـ منهجية البحث Research Methodology

ان الغرض من كتابة منهجية البحث هو لوصف طبيعة تصميم البحث والعينة اذا لم يتم دراسة مجتمع الدراسة وطرق جمع البيانات.

لذلك، فان الهدف من كتابة منهجية البحث هو ما يلي :

١ ـ وصف طبيعة تصميم البحث. Nature of the research design

٢ ـ طريقة اختيار العينة. Sampling Plan

٣ ـ طريقة جمع البيانات. Data Collection Method

٤ ـ اجراءات التحليل. Analysis Procedures

ان كتابة هذا الجزء من متن البحث ليست بالمهمة السهلة. ان هذا الجزء يجب أن يتضمن شرح كافي لما يجب ايصاله الى الشخص القارىء حتى يستطيع تقدير (Appreciate) طبيعة المنهجية المستخدمة. وعليه، فان عرض هذا الجزء يجب أن لا يكون ممل (boring) ولا يكون صعب جداً (overpowering).

يجب إن يبين جزء المنهجية للشخص القارىء ما اذا كان التصميم تمهيدي (استكشافي) (Exploratory) أو نهائي (Conclusive). كما يجب أن يبين ان هذا الجزء وبالتفصيل مصادر البيانات الأولية والثانوية. هذا بالاضافة الى تحديد طرق جمع البيانات وبدقة وطبيعتها. كما ان هذا الجزء يجب أن يهتم بالتوضيح للشخص القارئ عن الأفراد الذين تضمنتهم عينة الدراسة وحجمها وطبيعة اجراءات اختيار العينة.

لهذا، فان جزء المنهجية مصمّم لتحقيق ما يلي :

١ ـ تلخيص الجوانب الفنية (Technical Aspects) في مشروع البحث بطريقة (Style) تكون قابلة للفهم (Comprehensible) من قبل الأشخاص غير الفنيين.

٢ ـ بناء الثقة (Confidence) في نوعية الاجراءات المستخدمة.

د ـ النتائج Results

غالباً ما يعطى الجزء الأكبر من تقرير البحث الى نتائج عملية البحث، والتي يجب أن تكون فيه مربوطة مع أهداف الدراسة والمعلومات المطلوبة. لهذا، يجب أن تكون طريقة عرض النتائج طريقة منطقية في عملية كشف المعلومات، وكأن الشخص الباحث يقوم باخبار قصة. ان طريقة عرض النتائج يجب أن لا تكون عرض لسلسلة من الجداول غير المنتهية، بل يجب أن تركز طريقة العرض على

نقطة محددة في كل مرحلة ومن ثم ربطها مع البحث ككل بطريقة منطقية. بمعنى آخر، ان عرض النتائج يتطلب تدفق منطقي للمعلومات حتى تستطيع تحقيق أغراض الأشخاص متخذي القرارات.

هـ- محددات البحث Research Limitations

ان كل مشروع بحث، ومهما بلغت مهارة الشخص الباحث، يتميز بوجود بعض نقاط الضعف والتي تحتاج الى ايصالها لجمهور القراء بطريقة واضحة ودقيقة (Clear and Concise manner). وفي هذه العملية، يجب أن لا يقوم الشخص الباحث بمهاجمة نقاط ضعف الدراسة الفرعية. إن الغرض من عرض محددات الدراسة هو ليس للتقليل من نوعية مشروع البحث، انما لاعطاء الشخص القارىء الفرصة للحكم (Judge) على صلاحية (Validity) نتائج مشروع البحث.

ان محددات مشاريع الأبحاث الادارية غالباً ما تتعلق بالاجراءات العينية والاستجابة من جمهور الدراسة وضعف المنهجية. وعليه، فان كتابة جزء الاستنتاجات والتوصيات سوف يتاثر بمحددات البحث المعترف بها والمدركة (Recognized and Aknowledged Limitation). ان توضيح محددات الدراسة للشخص القارىء غالباً ما تكون مسؤولية الشخص الباحث.

٧- الخلاصة والتوصيات. Conclusions and Recommendations

يجب أن تكون خلاصة أي مشروع بحث وتوصياته تسلسل منطقي لعرض نتائجه. فالخلاصة يجب أن تربط (Link) وبوضوح ما بين نتائج مشروع البحث (Research Finding) والمعلومات المطلوبة (Information Nedded)، وبالاعتماد على هذا الربط يجب تشكيل (Formulate) التوصيات للعمل.

ان كثير من المدراء والباحثين يشعروا بعدم وجوب كتابة التوصيات من قبل الشخص الباحث. ان هؤلاء المدراء يجادلوا (Argue) بأن توصيات العمل يجب ن تعكس خبرة متخذي القرارات عند التعامل مع نتائج مشروع البحث. وبما ان القليل من الباحثين الذين يمتلكون الخبرة في التعامل مع نتائج البحث عند عملية اتخاذ القرار، فان توصيات الشخص الباحث سوف توزن وبشكل كبير لمصلحة نتائج البحث.

وبالمقابل، فان الكثير من المدراء الآخرين يشعرون وبقوة بضرورة كتابة توصيات البحث من خلال الشخص الباحث. ان هؤلاء المجموعة يجادلون بالمنطق التالي : ما دام أن متخذي القرارات يدركون البيئة التي تم فيها عمل التوصيات، فانه سوف يكون هناك فوائد كثيرة لعرض التوصيات في تقرير البحث للشخص متخذ القرار.

٨ ـ المراجع References

يهدف هذا الجزء من التقرير الى عرض جميع الكتب والمجلات والدوريات والتقارير والوثائق وغيرها من المراجع التي كان قد اعتمد عليها الشخص الباحث في جمع معلوماته وتحليل بياناته وتفسير نتائجه.

٩ ـ الملحقات Appendixes

ان الغرض من وضع الملحقات في تقرير البحث هو لتوفير مكان لأي مادة كان ليس من الضروري وضعها في متن تقرير البحث. ويمكن أن تشمل هذه المادة الجداول والرسومات والقوانين وأشكال طرق جمع البيانات وتفصيل خطة اختيار العينة وتقدير الأخطاء الاحصائية وغير ذلك من المعلومات. وغالباً ما تتميز مادة الملحقات بما يلي :

١ ـ صعوبة وتخصص هذه المادة مقارنة مع المادة المعروضة بالتقرير الرئيسي.

٢ ـ تصميم هذه المعلومات في بعض الأحيان يكون لتحقيق رغبات الأشخاص الموجهين فنياً.

ـ عرض البيانات Presentation of Data

ان التخطيط المسبق لكيفية العرض للمعلومات الكثيرة، او لوصف الاجراءات الفنية في عملية البحث أمر مهم للغاية ولا يجوز تأجيل التفكير به لما له من فائدة كبيرة في عملية ايصال المعلومات بسرعة الى جمهور القراء البحث. ان اكثر الأشكال المساعدة والمستخدمة في تقرير الأبحاث هي الجداول والأشكال (Tables and Graphs). فبالاضافة الى ان هذه الجداول تجعل قراءة تقرير البحث اكثر سهولة، فانها تعمل على تحسين المظهر بالنسبة لمتن التقرير. لذلك، فاننا نركز في هذا الجزء على الطرق المختلفة التي تستخدم لعرض البيانات الكمية بواسطة الجداول والأشكال.

ـ خطوط عامة لعرض الأشكال المساعدة

General Guidelines for Presenting Graphic Aids

انه لمن المفيد جداً وضع بعض التوضيحات في متن التقرير، اذا كان القارىء بحاجة للرجوع اليها اثناء قراءة تقرير البحث. كما يجب ان توضع الأشكال التوضيحية بالقرب من المادة المتعلقة بها كلما كان ذلك ممكناً. واذا كانت المعلومات الموضحة بجدول او شكل هي معلومات مكملة او طويلة جداً، فانه يمكن وضع هذه المعلومات في ملحق.

كما يفضل دائماً ان يتم عرض التوضيح للبيانات قبل عرضها بشكل مفصل، مع ذكر عدد قليل من الجمل التي تبرز الجوانب المهمة في البيانات والضرورية لتطوير البحث. كما يجب ان تتضمن الأشكال العناصر (Elements) التالية :

١ ـ رقم الجدول أو الشكل (Table or Figure Number). ان وضع الرقم او الجدول يسهل عملية الحصول على الجدول او الشكل ذو الاهتمام من قبل الشخص الباحث.

٢ ـ العنوان (Title). يجب أن يشير (Indicate) عنوان الجدول أو الشكل الى المحتوى وبوضوح.

٣ ـ الهامش (الحاشية) (Footnote). ان الهدف من كتابة الحاشية هو اما لتوضيح (Explain) أو تأهيل (Qualify) بند أو جزء من الجدول أو الشكل.

بشكل عام، ان البيانات تعرض اما من خلال الجداول أو الأشكال. وتعتبر الجداول طريقة منظمة لعرض البيانات الرقمية أو العددية اما من خلال أعمدة رأسية أو صفوف أفقية حسب الفئات المستخدمة في تصنيف البيانات. والجدول رقم (١٤ ـ ١) يعتبر توضيح لذلك، علماً بأن البيانات الواردة في هذا الجدول تتعلق بمعرفة عدد الأشخاص اللذين يزورون الأفرع المختلفة لسوبر ماركت في مدينة عمان خلال شهر آب من عام ١٩٩٥.

بينما الأشكال هي أداة العرض البيانات من خلال الرسم، وذلك من خلال اعطاء حجم معين من الشكل لكل مجموعة من البيانات المتشابهة. وأهم ما يميز التمثيل بواسطة الأشكال هو السهولة (Simplicity). لذلك، تعتبر الجدوال والأشكال طرق سريعة (Quik) وجذابة (Attractive) في عملية ايصال (Communicate) أرقام (Number) أو اتجاهات (Trends) أو علاقات (Relationships).

جدول رقم (١٤ - ١)

عدد الزائرون الشهري للمواقع المختلفة للسوبر ماركت

(شهر آب ، ١٩٩٥)

الفرع	عدد الزوار الداخلون	النسبة من الكل
الغربي	١٥٠	٪١٥
الشمالي	٢٠٠	٪٢٠
الجنوبي	٢٨٠	٪٢٨
الشرقي	٢٤٠	٪٢٤
الوسط	١٣٠	٪١٣
المجموع	١٠٠٠	٪١٠٠

وبعد أن يتم الشخص الحصول على البيانات المطلوبة، فانه يتوجب عليه أن يبدأ عملية البحث عن أفضل طريقة لعرض البيانات. ان مهمة البحث عن أفضل طريقة لعرض البيانات قد تبدو سهلة للحظة الأولى، الا ان اتخاذ القرار المتعلق بعملية اختيار الطريقة الأكثر ملائمة لعرض البيانات قد تكون مهمة صعبة (Arduous Task) . ان أكثر الأشكال استخداماً لعرض البيانات الادارية هي :

١ - القطع الدائرية . Pie Chart

٢ - الأعمدة البيانية . Bar Chart

٣ - الخطوط البيانية . Line Chart

ان سبب الاستخدام الشائع لهذه الطرق في تمثيل البيانات الادارية هو مقدرة هذه الطرق في تمثيل البيانات المعقدة (Complex Data) بطريقة واضحة وسهلة ومرئية

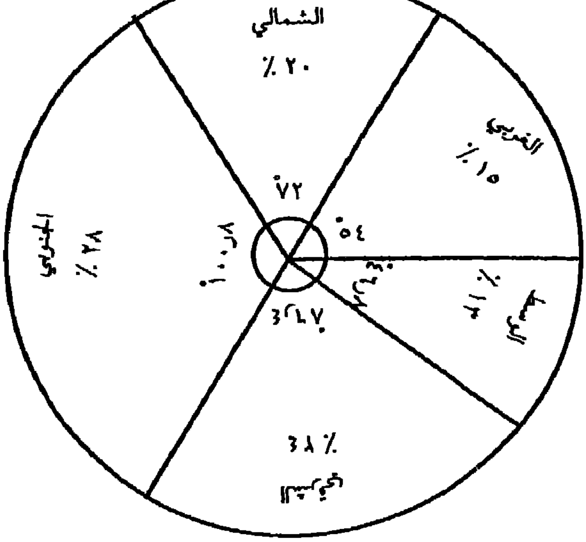
ومباشرة (Direct Visual Representation). اما الآنو فاننا سوف نقوم بشرح هذه الطرق بشيء من التفصيل.

١ ـ القطع (الأشكال) الدائرية Pie Chart

تعتبر الأشكال أو القطع الدائرية أداة لعرض مجموعة من البيانات بواسطة الرسم. وتعتبر القطع الدائرية من أكثر الطرق سهولة (Simplest) وفعالية (Effective) في تمثيل البيانات من أجل اظهار النسب المختلفة لمجموعات البيانات المختلفة. كما ان استخدام هذه الأشكال بعناية ودقة كافية قد يعبر عن بعض الخصائص أو الجوانب من البيانات بطريقة مرئية وواضحة وتسهل على القارىء فهم هذه البيانات. وبالغالب لا يكفي الشكل أو الجدول عن النص اللغوي، وانما يتم استخدامه للتأكيد على بعض العلاقات التي قد يراها الباحث مهمة. ويقوم مبدأ عمل هذه الطريقة في تمثيل البيانات عن طريق تقسيم الدائرة الى عدد من القطاعات أو الأجزاء حسب عدد خصائص العينات المراد مقارنتها أو معرفة العلاقة فيما بينها، بحيث يعكس حجم كل قطاع نسبة البيانات التي يمثلها ذلك القطاع مقارنة مع البيانات الكلية. ويمثل الشكل رقم (١٤ ـ ١) عرضاً للبيانات الواردة في الجدول رقم (١٤ ـ ١). لاحظ كيف ان حجم كل قطاع من القطاعات فعال (Active) في توضيح الفرق في نسبة عدد الزائرين لكل فرع مقارنة مع عدد الزائرين الكلي لمجموع الزائرين.

الشكل رقم (١٤ ـ ١)

القطاعات الدائرية لعدد الزائرين في شهر آب، ١٩٩٥م

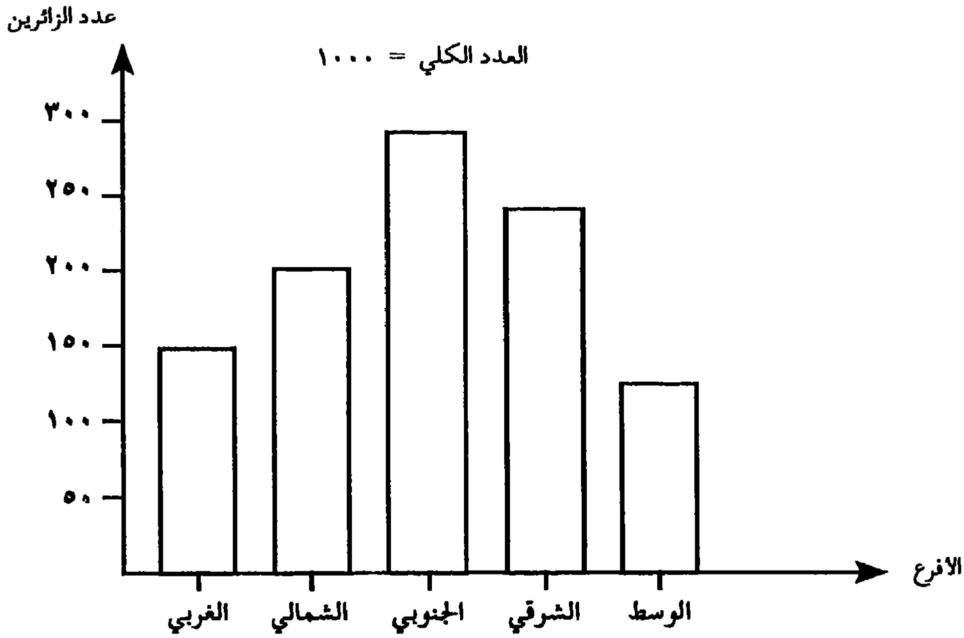


٢ ـ الأعمدة البيانية Bar Chart

توضح الاعمدة البيانية حجم البيانات التي يمثلها كل عمود من خلال مقارنة اطوال الأعمدة المرسومة مع بعضها البعض. واذا ما تم تصميم الأعمدة وبدقة، فان هذا سوف يؤدي الى فهم ما تعبر عنه البيانات التي تم تمثيلها بالأعمدة وبسهولة. كما تعتبر الأعمدة البيانية من أكثر الطرق استخدام في الحياة العملية. ويمثل الشكل رقم (١٤ ـ ٢) عرضاً للبيانات ذات العلاقة بالسوبر ماركت باستخدام الأعمدة البيانية وغالباً ما تستخدم الأعمدة البيانية لتوضيح علاقات المقارنة المتعددة (Multiple Comparison) والمعقدة (Complex). ولمقارنة مجموعات مختلفة من البيانات في شكل واحد، فانه يمكن استخدام شكل الأعمدة البيانية العنقودية (Cluster Bar Graph).

شكل رقم (١٤ ـ ٢)

الأعمدة البيانية لعدد الزائرين في شهر آب، ١٩٩٥



٣ ـ الخطوط البيانية Line Charts

غالباً ما تكون الخطوط البيانية فعالة في تفسير الأنماط (Trends) خلال فترة زمنية معينة. ان هذا النوع من الأشكال يستخدم الخطوط المتصلة (Continuous Line) لمتابعة أو ملاحقة (Trace) العلاقة ما بين نقاط البيانات. والبيانات الممثلة في الشكل رقم (١٤ ـ ٢) يمكن تمثيلها على شكل بياني. والشكل رقم (١٤ ـ ٣) يمثل الخط البياني لتفسير الفروق في عدد الزوار لكل فرع من فروع السوبر ماركت خلال شهر آب، ١٩٩٥م.

ويفضل استخدام الخطوط البيانية بدلاً من القطع الدائرية في الحالات التالية :

١ ـ اذا كانت البيانات تتضمن فترة طويلة (Long Time Period) .

٢ ـ في حالة تمثيل مجموعة مختلفة من البيانات على نفس الشكل.

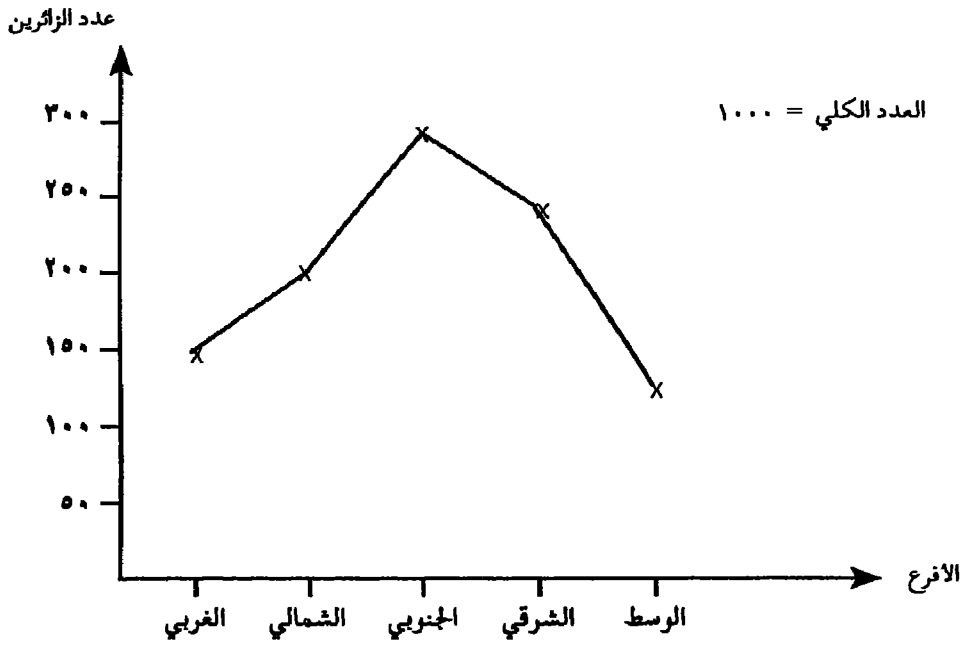
٣ ـ في الحالات التي يكون فيها التركيز على حركة البيانات أكثر من التركيز على البيانات نفسها.

٤ ـ في الحالات التي يكون فيها مطلوب عرض التوزيع التكراري (Frequency Distribution) للنمط أو الاتجاه.

٥ ـ في الحالات التي يكون فيها مطلوب اظهار التقدير (Estimation) أو التنبؤ (Forecaste) أو الاستقراء (Extrapolation) .

شكل رقم (١٤ ـ ٣)

الخط البياني لعدد الزائرين في شهر آب، ١٩٩٥م



ـ عرض الأبحاث الشفوي Oral Presentation

في كثير من الأحيان، نجد ان بعض المؤسسات أو الشركات أو الجهات التي تقوم بدعم البحث تقوم على طلب عرض البحث شفوياً من الشخص الباحث، فاذا ما كان مطلوب من الشخص الباحث القيام بعرض نتائج البحث شفوياً، فانه لا بد ان يتأكد من الأمور التالية قبل البدء بعملية العرض.

١ ـ التأكد من جميع التجهيزات (Equipment) والتي يمكن ان تستخدم في عملية العرض مثل الاضاءة وجهاز الصوت واداة تسليط الضوء (Projector) أو اي اداة مساعدة.

٢ ـ العمل على امتلاك خطة طوارىء (Contingency Plan) لمعالجة اي خلل طارىء لأي اداة في اقل وقت ممكن.

٣ ـ تحليل الجمهور (Audience Analysis). يمكن تحليل الجمهور من خلال التوقع لكيفية تفاعلهم مع نتائج الدراسة. هل سيكون الجمهور مع النتائج أو معارض أم محايد. وبناءاً على توقع ردة فعل الجمهور مع نتائج الدراسة، فانه يتم تحديد جمل الافتتاح (Opening Statements) للعرض الشفوي لنتائج الدراسة. كما انه من الحكمة (Wisdom) ان يبدأ الشخص الباحث العرض بتلك الأفكار (Ideas) التي يتوقع أنها تتفق مع الجمهور.

٤ ـ التدريب على ممارسة العرض أكثر من مرة قبل العرض الحقيقي (Practice Presentation Several Times). كما يفضل أن تتم الممارسة امام مجموعة من الأشخاص، اذا كان بالامكان، من أجل الحصول على تغذية عكسية (Feedback) والتي من شأنها ان تحسن (Improve) فعالية (Effectiveness) العرض عند الشخص الباحث.

٥ ـ العمل على بدء (Start) العرض بملخص عام (Overview) عن الدراسة. بمعنى آخر، يفضل أن يبدأ الشخص الباحث العرض عن طريق اخبار الجمهور بالمواضيع

الرئيسية التي سوف يقوم بعرضها.

٦ ـ العمل على مواجهة الجمهور في جميع الأوقات. (Face the Audience at all Times)

٧ ـ العمل على التحدث مع الجمهور أو متخذ القرار بدلاً من القراءة عن نص مكتوب (Script) أو شاشة جهاز العرض. بمعنى آخر، العمل على استخدام الملاحظات فقط لضمان عدم اسقاط أي نقاط رئيسية والمحافظة على متابعة عملية العرض بطريقة منظمة (Organized Manner).

٨ ـ استخدام وسائل المساعدة البصرية (Visual Aids) بفعالية. ان الجداول والأشكال المستخدمة يجب أن تكون بسيطة وسهلة القراءة. بمعنى آخر، يجب أن يؤدي استخدام وسائل المساعدة الى تبسيط أو تسهيل عملية العرض وليس العكس.

٩ ـ تجنب الأسلوب الملهي خلال الحديث (Avoid Distracting mannerisms while Speaking). ويتم تجنب هذا الأسلوب من خلال التأكد من أن كل حركة يقوم بها الشخص الباحث تحقق غرضاً معيناً. كما يجب على الشخص الباحث أن يقلل من استخدام بعض الكلمات مثل «كما تعلمون»، «أم»، «نعم» خلال الكلمات أو الجمل.

١٠ ـ تذكير الجمهور (Reminding Audience) بما اذا كان هناك أي سؤال بعد الانتهاء من عرض البحث. كما يجب على الشخص الباحث أخذ الأمور التالية بعين الاعتبار اثناء طرح الأسئلة :

أ ـ التركيز (Concentrate) على السؤال. بمعنى آخر، لا يجوز التفكير بالاجابة الا بعد الانتهاء من طرح السؤال.

ب ـ اعادة السؤال (Repeat the Question). اذا كان السؤال صعباً (Tough)، فيجب العمل على محاولة اعادة صياغة (Rephrase) السؤال. ان الهدف

من اعـادة الصـيـاغـة هو لضـمـان فـهـم السـؤال من قـبل كل فـرد من افـراد الجمهور. هذا بالاضافة الى ان اعادة تشكيل السؤال يعطي الشخص الباحث الوقت الكافي لتشكيل (Formulate) الاجابة على السؤال.

جـ ـ لا تبـتـذل (تبـهـت) الاجـابة (Don't Fake an Answer). يجب على الشخص الباحث الاعتراف (Admit) بعدم معرفة الاجابة اذا كان السؤال صعباً، ومن ثم اخبار الشخص السائل بانك سوف تحاول الحصول على اجابة. وبعد الانتهاء من عرض تقرير البحث، يجب العمل على معرفة مكان وجود الشخص السائل من اجل تزويده بالاجابة على السؤال حين الحصول عليها.

د ـ العمل على اجابة الأسئلة باختصار (Brief) ومن ثم تدعيم الاجابة بالدليل اذا كان ذلك ممكناً.

ـ النواحي الفنية في كتابة تقرير البحث.

ان كتابة تقرير البحث تتطلب بالاضافة الى مقدرة الشخص الباحث على جمع البيانات وتحليلها وتفسير نتائجها، تحتاج الى الاخذ بعين الاعتبار المسائل الفنية في ترتيب المحتويات وتوثيق المصادر واعداد قائمة المراجع وغير ذلك من المسائل الفنية. لذلك، فاننا سوف نتعرض في هذا الجزء الى شرح بعض الجوانب الفنية في كتابة تقرير البحث والتي منها التوثيق والاقتباس والحواشي وقائمة المراجع.

١ ـ التوثيق

ان البحث العلمي عبارة عن جهد انساني مستمر ومتصل بحيث يستلزم من الشخص الباحث مسح (Survey) جميع الجهود السابقة والاضافة عليها والتمهيد لباحثين في المستقبل. لذلك، فاننا كثير من نجد ان بعض الباحثين يشيروا الى نتائج

غيرهم ويستخدموها ويبنوا عليها في بعض الأحيان، أو ينتقدوها ويظهروا جوانب العجز فيها. وقد تلزم عملية البحث في بعض الأحيان الى الاستفادة من أفكار غيرهم واستخدامها اما بنصها الحرفي أو اعادة صياغتها بشكل يتلائم مع مشكلة البحث قيد الدراسة. لهذا، فان أعراف البحث ومبادئه تقتضي الاشارة الى المصادر التي استخدمها الشخص الباحث وتوثيقها. وعملية التوثيق هذه لا تعبر عن الأمانة العلمية فحسب، بل تعبر عن الطريقة التي تتفاعل فيها المعرفة العلمية لدى أجيال الباحثين، وكذلك المعيار الذي يستخدم في الحكم على مدى مساهمة البحث في تقدم المعرفة واستمرار نموها. ويوجد هناك طرق مختلفة للتوثيق منها ما يلي :

أ ـ التوثيق في متن التقرير.

ان المقصود بالتوثيق في متن التقرير هو الاشارة الى المرجع اثناء كتابة نص متن التقرير.

ب ـ التوثيق في قائمة المراجع.

ان المقصود بالتوثيق في قائمة المراجع هو الاشارة الى جميع المراجع التي تم استخدامها فعلاً في متن التقريرفي قائمة المراجع. كما يشترط أن تكون هذه المراجع منشورة ويمكن الحصول عليها أو الرجوع بسهولة ويسر. وغالباً ما ترتب المراجع أبجدياً حسب الأحرف الأبجدية للحروف الأولى من أسماء المؤلفين، أو اعطاءها أرقاماً متسلسلة داخل متن البحث ومن ثم ترتيبها حسب تسلسل الأرقام. واذا استخدم الباحث أكثر من نوع من المراجع، فانه يجب كتابة المراجع العربية في قائمة والأجنبية في قائمة أخرى. وغالباً ما تستخدم العناصر التالية كعناصر توثيق في قائمة المراجع والتي هي :

١ ـ اسم المؤلف ٢ ـ سنة النشر ٣ ـ عنوان المرجع ٤ ـ عنوان دار النشر.

٢ ـ الاقتباس

ان المقصود بالاقتباس هو الاستفادة من الأفكار التي وردت في دراسات الباحثين الآخرين. ويعتبر الاقتباس من الأمور المهمة والضرورية والتي لا يستطيع الباحث الاستغناء عنها من أجل اكمال البحث قيد الدراسة. لذلك، يجب على الشخص الباحث أن يعطي الاقتباس اهمية خاصة من حيث الدقة واختيار الاقتباس المناسب والمصدر المناسب.

كما يمكن تقسيم الاقتباس الى نوعين رئيسين هما :

أ ـ الاقتباس الحرفي (المباشر). وهو ذلك النوع من الاقتباس الذي يقوم به الشخص الباحث عند نقل نصاً حرفياً بالتمام والكمال وكما ورد في المصدر الأصلي.

ب ـ الاقتباس غير الحرفي (غير المباشر). وهو ذلك النوع من الاقتباس الذي يقوم به الشخص الباحث عند الاستفادة من فكرة معينة موجودة في مصدر بحيث يقوم بصياغتها من جديد بشكل تتلائم مع مادة البحث.

٣ ـ الحواشي

ان عملية البحث تحتاج في كثير من الأحيان الى الاستعانة بآراء وأفكار الباحثين الآخرين. وعليه، فان الأمانة العلمية تستلزم من الشخص الباحث الاشارة الى المصدر الذي استقى منه الباحث الرأي أو الفكرة اعترافاً وتقديراً لما بذله ذلك الباحث في اعداد مادة البحث. هذا بالاضافة الى وجوب الاشارة الى أي مصدر آخر كان الباحث قد استخدمه عند اعداد مادة بحثه. ويوجد هناك نوعين رئيسين للحواشي هما :

أ ـ حاشية المحتوى.

تستخدم حاشية المحتوى بالغالب الى توضيح فكرة معينة كانت قد وردت في متن البحث بشيء من التفصيل، حتى يتمكن الشخص الباحث من سرد أفكار في

متن البحث بشكل منطقي. وغالباً ما يستخدم الشخص الباحث الأرقام للاشارة الى المعلومات الواردة في الحواشي. ولكن هذا لا يمنع من استخدام أي شيء آخر عدا الأرقام للاشارة الى ما ظهر في الحواشي مثل الرموز أو الاشارات أو اي شيء آخر.

ب ـ حاشية المصدر.

غالباً ما تستخدم حاشية المصدر من أجل اظهار اسم المصدر الذي استقى منه الباحث معلوماته، أو لإرجاع الشخص القارىء الى مصادر أخرى في حالة وجود الرغبة عند الشخص القارىء في الاستزادة أو التوسع في موضوع معين.

٤ ـ قائمة المراجع

تتضمن قائمة المراجع جميع المصادر التي استخدمها الشخص الباحث واعتمد عليها في اعداد البحث. وتشمل المراجع على أي شيء له علاقة بموضوع البحث مثل الكتب والمجلات والدراسات والوثائق الحكومية وغير الحكومية والمحاضرات واي مصدر آخر. وغالباً ما توضع قائمة المراجع بالبحث بعد المتن مباشرة وقبل الملحقات. وتختلف طريقة كتابة المراجع بالنسبة للكتب عنه بالنسبة للدوريات. ففي حالة الكتب، يجب أن يوضع خط تحت اسم الكتاب للتعبير عن أن هذا المصدر هو كتاب وكما يلي :

جمال زكي والسيد يس ، أسس البحث الاجتماعي ، دار الفكر العربي، ١٩٦٢ .

أما في حالة كتابة الدوريات، فانه يجب ان يوضع الخط تحت اسم الدورية لتمييز الدورية عن الكتاب وكما يلي :

محمد عليان ، تاثير الشراء بالوقت المحدد على تكاليف النقل والمخزون، إدارة اليابان ، المجلد رقم (٩)، العدد رقم (٣)، ١٩٨٣ .

أسئلة

س١ : لماذا يعتبر تقرير البحث مهماً ؟

س٢ : ناقش الخطوط الاسترشادية عند كتابة تقرير البحث ؟

س٣ : كيف يمكن للعرض الشفهي لنتائج البحث ان يكون مكملاً لتقرير البحث المكتوب ؟

س٤ : أذكر عناصر متن البحث ؟

س٥ : هل تعتقد انه من الضروري كتابة ملخص الدراسة، ولماذا ؟

س٦ : ناقش طرق عرض البيانات ؟

س٧ : ما هو المقصود بالاقتباس، وما هي أنواعه ؟

س٨ : كيف يمكن ان تميز فيما اذا كان المرجع كتاب أم دورية ؟

الملاحق

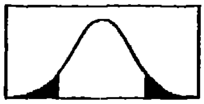
جدول (١)

الأرقام العشوائية

10 09 73 25 33	76 52 01 35 35	34 67 35 48 76	80 95 90 91 17	39 29 27 49 45
37 54 20 48 05	64 89 47 42 96	24 80 52 40 37	20 63 61 04 02	00 82 29 16 65
08 42 26 89 53	19 64 50 93 03	23 20 90 25 60	15 95 33 47 64	35 08 03 36 06
90 01 90 25 29	09 37 67 07 15	38 31 13 11 65	88 67 67 43 97	04 43 62 76 59
12 80 79 99 70	80 15 73 61 47	64 03 23 66 53	98 95 11 68 77	12 17 17 68 33
66 06 57 47 17	34 07 27 08 50	36 69 73 61 70	65 81 33 98 85	11 19 92 91 70
31 06 01 08 05	45 57 18 24 06	35 30 34 26 14	86 79 90 74 39	23 40 30 97 32
85 26 97 76 02	02 05 16 56 92	68 66 57 48 18	73 05 38 52 47	18 62 38 85 79
63 57 33 21 35	05 32 54 70 48	90 55 35 75 48	28 46 82 87 09	83 49 12 55 24
73 79 64 57 53	03 52 96 47 78	35 80 83 42 82	60 93 52 03 44	35 27 38 84 35
98 52 01 77 67	14 90 56 86 07	22 10 94 05 58	60 97 09 34 33	50 50 07 39 98
11 80 50 54 31	39 80 82 77 32	50 72 56 82 48	29 40 52 42 01	52 77 56 78 51
83 45 29 96 34	06 28 89 80 83	13 74 67 00 78	18 47 54 06 10	68 71 17 78 17
88 68 54 02 00	86 50 75 84 01	36 76 66 79 51	90 36 47 64 93	29 60 91 10 62
99 59 46 73 48	87 51 76 49 69	91 82 60 89 28	93 78 56 13 68	23 47 83 41 13
65 48 11 76 74	17 46 85 09 50	58 04 77 69 74	73 03 95 71 86	40 21 81 65 44
80 12 43 56 35	17 72 70 80 15	45 31 82 23 74	21 11 57 82 53	14 38 55 37 63
74 35 09 98 17	77 40 27 72 14	43 23 60 02 10	45 52 16 42 37	96 28 60 26 55
69 91 62 68 03	66 25 22 91 48	36 93 68 72 03	76 62 11 39 90	94 40 05 64 18
09 90 32 05 05	14 22 56 85 14	46 42 75 67 88	96 29 77 88 22	54 38 21 45 98
91 49 91 45 23	68 47 92 76 86	46 16 28 35 54	94 75 08 99 23	37 08 92 00 48
80 33 69 45 98	26 94 03 08 58	70 29 73 41 35	53 14 03 33 40	42 05 08 23 41
44 10 48 19 49	85 15 74 79 54	32 97 92 65 75	57 60 04 08 81	22 22 20 64 13
12 55 07 37 42	11 10 00 20 40	12 86 07 46 97	96 64 48 94 39	28 70 72 58 15
63 60 64 93 29	16 50 53 44 84	40 21 95 25 63	43 65 17 70 82	07 20 73 17 90
61 19 69 04 46	26 45 74 77 74	51 92 43 37 29	65 39 45 95 93	42 58 26 05 27
15 47 44 52 66	95 27 07 99 53	59 36 78 38 48	82 39 61 01 18	33 21 15 94 66
94 55 72 85 73	67 89 75 43 87	54 62 24 44 31	91 19 04 25 92	92 92 74 59 73
42 48 11 62 13	97 34 40 87 21	16 86 84 87 67	03 07 11 20 59	25 70 14 66 70
23 52 37 83 17	73 20 88 98 37	68 93 59 14 16	26 25 22 96 63	05 52 28 25 62
04 49 35 24 94	75 24 63 38 24	45 86 25 10 25	61 96 27 93 35	65 33 71 24 72
00 54 99 76 54	84 05 18 81 59	96 11 96 38 96	54 69 28 23 91	23 28 72 95 29
35 96 31 53 07	26 89 80 93 54	33 35 13 54 62	77 97 45 00 24	90 10 33 93 33
59 80 80 83 91	45 42 72 68 42	83 60 94 97 00	13 02 12 48 92	78 56 52 01 06
46 05 88 52 36	01 39 09 22 86	77 28 14 40 77	93 91 08 36 47	70 61 74 29 41
32 17 90 05 97	87 37 92 52 41	05 56 70 70 07	86 74 31 71 57	85 39 41 18 38
69 23 48 14 06	20 11 74 52 04	15 95 66 00 00	18 74 39 24 23	97 11 89 63 38
19 56 54 14 30	01 75 87 53 79	40 41 92 15 85	66 67 43 68 06	84 96 28 52 07
45 15 51 49 38	19 47 60 72 46	43 66 79 45 43	59 04 79 00 33	20 82 66 95 41
94 86 43 19 94	36 16 81 08 51	34 88 88 15 53	01 54 03 54 56	05 01 45 11 76
93 08 62 48 26	45 24 02 84 04	44 99 90 88 96	39 09 47 34 07	35 44 13 18 80
33 18 51 62 32	41 94 15 09 49	89 43 54 85 81	88 69 54 19 94	37 54 87 30 43
80 95 10 04 06	96 38 27 07 74	20 15 12 33 87	25 01 62 52 98	94 62 46 11 71
79 75 24 91 40	71 96 12 82 96	69 86 10 25 91	74 85 22 05 39	00 38 75 95 79
18 63 33 25 37	98 14 50 65 71	31 01 02 46 74	05 45 56 14 27	77 93 89 19 36
74 02 94 39 02	77 55 73 22 70	97 79 01 71 19	52 52 75 80 21	80 81 45 17 48
54 17 84 56 11	80 99 33 71 43	05 33 51 29 69	56 12 71 92 55	36 04 09 03 24
11 66 44 98 83	52 07 98 48 27	59 38 17 15 39	09 97 33 34 40	88 46 12 33 56
48 32 47 79 28	81 24 96 47 10	02 29 53 68 70	32 30 75 75 46	15 02 00 99 94
69 07 49 41 38	87 63 79 19 76	35 58 40 44 01	10 51 82 16 15	01 84 87 69 38

جدول (٢)

جدول التوزيع الطبيعي ذو الطرفين

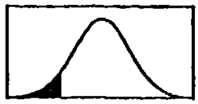
This table shows: 

x/σ	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0 0	1 0000	9920	9840	9761	.9681	.9601	.9522	.9442	.9362	.9283
0 1	9203	9124	.9045	.8966	.8887	.8808	.8729	.8650	.8572	.8493
0.2	.8415	.8337	8259	8181	.8103	.8026	.7949	.7872	.7795	.7718
0 3	.7612	.7566	.7490	7414	.7339	.7263	.7188	.7114	.7039	.6965
0 4	.6892	.6818	.6745	.6672	.6599	.6527	.6455	.6384	.6312	.6241
0.5	6171	.6101	.6031	.5961	.5892	.5823	.5755	.5687	.5619	.5552
0 6	.5485	5419	.5353	.5287	5222	.5157	.5093	5029	.4965	.4902
0 7	.4839	.4777	.4715	.4654	.4593	.4533	.4473	.4413	.4354	.4295
0.8	.4237	.4179	.4122	4065	.4009	.3953	.3898	.3843	.3789	.3735
0.9	.3681	.3628	.3576	.3524	.3472	.3421	.3371	.3320	.3271	.3222
1.0	.3173	3125	.3077	.3030	.2983	.2937	.2891	.2846	.2801	.2757
1 1	.2713	.2670	.2627	.2585	.2543	.2501	.2460	.2420	.2380	.2340
1 2	.2301	.2263	2225	.2187	.2150	.2113	.2077	.2041	.2005	.1971
1.3	.1936	.1902	.1868	.1835	.1802	.1770	.1738	.1707	.1676	.1645
1.4	.1615	1585	.1556	1527	1499	.1471	.1443	.1416	.1389	.1362
1 5	.1336	1310	.1285	1260	1236	.1211	.1188	.1164	.1141	.1118
1.6	.1096	1074	1052	.1031	.1010	.0989	.0969	.0949	.0930	.0910
1.7	0891	.0873	.0854	.0836	0819	0801	.0784	.0767	.0751	.0735
1 8	0719	.0703	.0688	0672	.0658	.0643	.0629	.0615	.0601	.0588
1.9	.0574	.0561	0519	.0536	.0524	0512	.0500	.0488	.0477	.0466
2.0	.0455	.0441	0434	0421	.0414	0404	.0394	.0385	.0375	.0366
2 1	.0357	.0349	0310	.0332	.0324	.0316	.0308	.0300	.0293	.0285
2.2	0278	.0271	0264	0257	.0251	.0244	.0238	.0232	.0226	.0220
2.3	0214	0209	0203	.0198	0193	0188	.0183	.0178	.0173	.0168
2 4	.0164	.0160	.0155	0151	.0147	.0143	.0139	.0135	.0131	.0128
2.4	0124	.0121	0117	0114	.0111	.0108	.0105	.0102	.00988	.00960
2 6	.00932	.00905	00879	00854	.00829	.00805	.00781	.00759	.00736	.00715
2 7	00693	00673	00653	00633	00614	.00596	.00578	.00561	.00544	.00527
2 8	.00511	00495	00180	00465	00451	.00437	.00424	.00410	.00398	.00385
2.9	.00373	00361	00350	.00339	.00328	.00318	.00308	.00298	.00288	.00279

x/σ	.0	.1	.2	.3	.4	.5	.6	.7	.8	.9
3	00270	.00191	00137	$.0^{4}967$	$.0^{3}674$	$.0^{3}465$	$.0^{3}318$	$.0^{3}216$	$.0^{3}145$	$.0^{4}962$
4	$.0^{4}633$	$.0^{4}413$	$0^{4}267$	$.0^{4}171$	$.0^{4}108$	$.0^{5}680$	$.0^{5}422$	$.0^{5}260$	$.0^{5}159$	$.0^{6}958$
5	$0^{6}573$	$0^{6}340$	$0^{6}199$	$0^{6}116$	$.0^{7}666$	$.0^{7}380$	$.0^{7}214$	$.0^{7}120$	$.0^{8}663$	$0^{8}364$
6	$0^{8}197$	$0^{8}106$	$.0^{9}565$	$.0^{9}298$	$.0^{9}155$	$.0^{10}803$	$.0^{10}411$	$.0^{10}208$	$.0^{10}105$	$.0^{11}520$

جدول (٣)

جدول التوزيع الطبيعي ذو الطرف

This table shows:  or 

x/σ	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.4960	.4920	4880	.4840	.4801	4761	.4721	.4681	.4641
0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	0294
1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
2.0	.0228	.0222	.0217	.0212	.0207	0202	.0197	.0192	.0188	.0183
2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
2.3	.0107	.0104	.0102	.00990	.00964	.00939	.00914	.00889	.00866	.00842
2.4	.00820	.00798	.00776	.00755	.00734	.00714	00695	.00676	.00657	00639
2.5	.00621	.00604	.00587	.00570	.00554	.00539	.00523	.00508	.00494	.00480
2.6	.00466	.00453	.00440	.00427	.00415	.00402	.00391	.00379	.00368	.00357
2.7	.00347	.00336	.00326	.00317	.00307	.00298	.00289	.00280	.00272	.00264
2.8	.00256	.00248	.00240	.00233	.00226	.00219	00212	.00205	.00199	.00193
2.9	.00187	.00181	.00175	.00169	.00164	.00159	.00154	.00149	.00144	.00139

x/σ	.0	.1	.2	.3	.4	.5	.6	.7	.8	.9
3	.00135	$.0^{3}968$	$.0^{3}687$	$.0^{3}483$	$.0^{3}337$	$.0^{3}233$	$.0^{3}159$	$.0^{3}108$	$.0^{4}723$	$0^{4}481$
4	$.0^{4}317$	$.0^{4}207$	$.0^{4}133$	$.0^{5}854$	$.0^{5}541$	$.0^{5}340$	$.0^{5}211$	$0^{5}130$	$.0^{6}793$	$.0^{6}479$
5	$.0^{6}287$	$.0^{6}170$	$.0^{7}996$	$.0^{7}579$	$.0^{7}333$	$.0^{7}190$	$.0^{7}107$	$.0^{8}599$	$.0^{8}332$	$.0^{8}182$
6	$.0^{9}987$	$.0^{9}430$	$.0^{9}282$	$.0^{9}149$	$.0^{10}777$	$.0^{10}402$	$.0^{10}206$	$.0^{10}104$	$.0^{11}523$	$.0^{11}260$

جدول (٤)

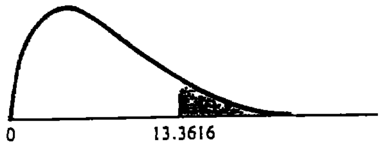
جدول ت ذو الطرف وذو الطرفين

TABLE A-4 TABLE OF CRITICAL VALUES OF *t*

Df	Level of significance for two-tailed test								Df
	0.5	0.4	0 3	0.2	0.1	0.05	0.02	0.01	
	Level of significance for one-tailed test								
	0.25	0 20	0 15	0 10	0.05	0.025	0.01	0.005	
1	1 000	1 376	1 963	3 078	6 314	12 706	31 821	63 657	1
2	816	1.061	1 386	1 886	2 920	4 303	6 965	9.925	2
3	.765	.978	1.250	1.638	2.353	3.182	4.541	5.841	3
4	.741	.941	1 190	1.533	2.132	2.776	3.747	4.604	4
5	727	.920	1.156	1 476	2.105	2 571	3.365	4.032	5
6	.718	906	1 134	1 440	1.943	2.447	3 143	3.707	6
7	711	896	1 119	1.415	1.895	2.365	2.998	3.499	7
8	706	.889	1 108	1 397	1.860	2 306	2.896	3.355	8
9	.703	.883	1 100	1 383	1.833	2.262	2.821	3.250	9
10	.700	.879	1 093	1.372	1.812	2 228	2.764	3.169	10
11	.697	.876	1 088	1 363	1.796	2 201	2.718	3.106	11
12	.695	.873	1 083	1 356	1.782	2.179	2.681	3.055	12
13	.694	.870	1 079	1 350	1.771	2.160	2.650	3.012	13
14	.692	.868	1 076	1 345	1 761	2 145	2.624	2.977	14
15	.691	.866	1 074	1 341	1.753	2.131	2 602	2.947	15
16	.690	.865	1 071	1 337	1.746	2.120	2.583	2 921	16
17	.689	.863	1 069	1.333	1.740	2 110	2.567	2 898	17
18	.688	.862	1 067	1 330	1.734	2 101	2 552	2.878	18
19	.688	861	1.066	1.328	1.729	2.093	2 539	2 861	19
20	687	.860	1.064	1.325	1.725	2.086	2.528	2 845	20
21	.686	.859	1 063	1.323	1.721	2.080	2.518	2.831	21
22	.686	.858	1.061	1.321	1.717	2.074	2.508	2.819	22
23	685	.858	1 060	1.319	1 714	2 069	2 500	2 807	23
24	.685	.857	1 059	1 318	1 711	2 064	2.492	2 797	24
25	684	856	1 058	1-316	1.708	2.060	2-485	2.787	25
26	684	856	1 058	1 315	1 706	2 056	2 479	2 779	26
27	.684	.855	1 057	1 314	1 703	2 052	2 473	2.771	27
28	683	.855	1 056	1.313	1 701	2.048	2.467	2 763	28
29	.683	854	1 055	1.311	1.699	2.045	2.462	2 756	29
30	.683	854	1 055	1 310	1.697	2.042	2.457	2 750	30
35						2 030		2 724	35
40						2 021		2 704	40
45						2.014		2.690	45
50						2 008		2 678	50
60						2 000		2 600	60
70						1 994		2.648	70
80						1.990		2-638	80
90						1-987		2 632	90
100						1.984		2 626	100
125						1.979		2 616	125
150						1 976		2.609	150
200						1 972		2.601	200
300						1 968		2 592	300
400						1 966		2 588	400
500						1 965		2.568	500
1000						1 962		2 581	1000
∞	67149	84162	1 03643	1 28155	1.64485	1 95996	2 32634	2 57582	∞

جدول (٥)

جدول توزيع كاي تربيع $(X)^{٢}$



Degrees of freedom	Probabilities 0.990	0.975	0.950	0.900	0.100	0.050	0.025	0.010
1	157088.10^{-9}	982069.10^{-9}	393214.10^{-8}	0.0157908	2.70554	3.84146	5.02389	6.63490
2	0.0201007	0.0506356	0.102587	0.210720	4.60517	5.99147	7.37776	9.21034
3	0.114832	0.215795	0.351846	0.584375	6.25139	7.81473	9.34840	11.3449
4	0.297110	0.484419	0.710721	1.063623	7.77944	9.48773	11.1433	13.2767
5	0.554300	0.831211	1.145476	1.61031	9.23635	11.0705	12.8325	15.0863
6	0.872085	1.237347	1.63539	2.20413	10.6446	12.5916	14.4494	16.8119
7	1.239043	1.68987	2.16735	2.83311	12.0170	14.0671	16.0128	18.4753
8	1.646482	2.17973	2.73264	3.48954	13.3616	15.5073	17.5346	20.0902
9	2.087912	2.70039	3.32511	4.16816	14.6837	16.9190	19.0228	21.6660
10	2.55821	3.24697	3.94030	4.86518	15.9871	18.3070	20.4831	23.2093
11	3.05347	3.81575	4.57481	5.57779	17.2750	19.6751	21.9200	24.7250
12	3.57056	4.40379	5.22603	6.30380	18.5494	21.0261	23.3367	26.2170
13	4.10691	5.00874	5.89186	7.04150	19.8119	22.3621	24.7356	27.6883
14	4.66043	5.62872	6.57063	7.78953	21.0642	23.6848	26.1190	29.1413

تابع / جدول (٥)

جدول توزيع كاي تربيع (X)٢

Degrees of freedom	Probabilities 0.990	0.975	0.950	0.900	0.100	0.050	0.025	0.010
15	5.22935	6.26214	7.26094	8.54675	22.3072	24.9958	27.4884	30.5779
16	5.81221	6.90766	7.96164	9.31223	23.5418	26.2962	28.8454	31.9999
17	6.40776	7.56418	8.67176	10.0852	24.7690	27.5871	30.1910	33.4087
18	7.01491	8.23075	9.39046	10.8649	25.9894	28.8693	31.5264	34.8053
19	7.63273	8.90655	10.1170	11.6509	27.2036	30.1435	32.8523	36.1908
20	8.26040	9.59083	10.8508	12.4426	28.4120	31.4104	34.1696	37.5662
21	8.89720	10.28293	11.5913	13.2396	29.6151	32.6705	35.4789	38.9321
22	9.54249	10.9823	12.3380	14.0415	30.8133	33.9244	36.7807	40.2894
23	10.19567	11.6885	13.0905	14.8479	32.0069	35.1725	38.0757	41.6384
24	10.8564	12.4011	13.8484	15.6587	33.1963	36.4151	39.3641	42.9798
25	11.5240	13.1197	14.6114	16.4734	34.3816	37.6525	40.6465	44.3141
26	12.1981	13.8439	15.3791	17.2919	35.5631	38.8852	41.9232	45.6117
27	12.8786	14.5733	16.1513	18.1138	36.7412	40.1133	43.1944	46.9630
28	13.5648	15.3079	16.9279	18.9392	37.9159	41.3372	44.4607	48.2782
29	14.2565	16.0471	17.7083	19.7677	39.0875	42.5569	45.7222	49.5879
30	14.9535	16.7908	18.4926	20.5992	40.2560	43.7729	46.9792	50.8922
40	22.1643	24.4331	26.5093	29.0505	51.8050	55.7585	59.3417	63.6907
50	29.7067	32.3574	34.7642	37.6886	63.1671	67.5048	71.4202	76.1539
60	37.4848	40.4817	43.1879	46.4589	74.3970	79.0819	83.2976	88.3794
70	45.4418	48.7576	51.7393	55.3290	85.5271	90.5312	95.0231	100.425
80	53.5400	57.1532	60.3915	64.2778	96.5782	101.879	106.629	112.329
90	61.7541	65.6466	69.1260	73.2912	107.565	113.145	118.136	124 116
100	70.0648	74 2219	77.9295	82.3581	118.498	124.342	129.561	135.807

المراجع

* المراجع العربية *

١ - أحمد بدر، أصول البحث العلمي ومفاهيمه، الكويت : وكالة المطبوعات، ١٩٧٨.

٢ - أحمد بدر، أصول البحث العلمي ومناهج، الطبعة الخامسة، القاهرة : دار المعارف، ١٩٨٩.

٣ - أحمد شلبي، كيف تكتب بحثاً أو رسالة ، الطبعة الثانية ، القاهرة : دار النهضة العربية، ١٩٧٤.

٤ - ثريا ملحس، منهج البحوث العلمية للطلبة الجامعيين ، بيروت : مكتبة المدرسة ودار الكتاب اللبناني، ١٩٦٠.

٥ - جمال زكي و السيد يس ، أسس البحث الاجتماعي، القاهرة : دار الفكر العربي، ١٩٦٢.

٦ - حلمي محمد وعبد الرحمن صالح، المرشد في كتابة الأبحاث، بيروت : دار الفكر، ١٩٧٥.

٧ - عبد الباسط محمد، أصول البحث الاجتماعي، القاهرة : مكتبة الانجلو المصرية، ١٩٧١.

٨ - عبد الله عبد الدايم، التربية التجريبية والبحث التربوي، بيروت : دار العلم للملايين، ١٩٨١.

٩ - محمد أبو صالح وعدنان عوض ، مقدمة في الاحصاء ، نيويورك : دار جون وايلي وأبنائه، ١٩٨٣.

١٠ - عمر الشيباني، مناهج البحث الاجتماعي، بيروت : دار الثقافة، ١٩٧١.

١١ - فوزي غرايبة وزملاءه ، أساليب البحث العلمي في العلوم الاجتماعية والانسانية ، عمان : ١٩٨٧.

١٢ ـ ڤان دالين، مناهج البحث في التربية وعلم النفس ، ترجمة محمد نبيل وآخرين ، القاهرة : مكتبة الانجلو مصرية، ١٩٧٩ .

١٣ ـ كايد عبد الحق، مبادىء في كتابة البحث العلمي، دمشق: مكتبة الفتح، ١٩٧٢ .

١٤ ـ محمد فتحي، الاحصاء في اتخاذ القرارات التجارية، القاهرة: مكتبة عين الشمس، ١٩٦٩.

١٥ ـ محمد مظلوم، طرق الاحصاء، الطبعة الخامسة، القاهرة : دار المعارف، ١٩٦٥ .

١٦ ـ محمد الجوهري و عبد الله الخريجي ، طرق البحث الاجتماعي ، القاهرة : دار الكتاب للتوزيع، ١٩٨٢ .

* المراجع الانجليزية *

1- Addams, R. and Presis, J., Human Organization Research, Homewood, Illinois, The Dorsey Press, Inc., 1960.

2- Boyed, H., and Ralf, A., Marketing Research : Text and Cases, Homewood, ill : Richard Irwin, 1956.

3- Campbell, C. and Joiner, B., How to Get the Answer Without Being Sure You Have Asked The Question, The American Statistician, Vol. 27, Dec., 1973.

4-Churchill, G., Marketing Research : Methodological Foundation, Hillsdale, ill : Dryden Press, 1976.

5-Donald, S., and Gerald, S., Survey Research : A Decisional Approach, New York : Intext Educational Publisher, 1973.

6- Galtung, J., Theory & Methods of Social Research, New York : Columbia University Press, 1967.

7- Kerlinger, F., Foundations of Behavioural Research, 2nd ed., New York : Holt, 1973.

8- Mansfield, E., Staristics for Business : Methods and Applications, 2nd ed. New York : Norton & Company, Inc., 1983.

9- Mason, R, Statistical Techniques in Business & Economics, Homewood, ill : Richard, D. Irwin, inc., 1967.

10- Neter, J, Wasserman, W. and Whitmore, G., Applied Statistics, 5th ed. Nre York : Allyn & Bacon, Inc., 1979.

11- Phillips, B., Social Research : Strategy & Tactics, New York : The Macmillan Company, 1966.

12- Reeder, W., How to Write a Thesis, Bloomington, ill, : Public School Publishing Company, 1925.

13- Robin, R., A Porcedure for Securing Returns to Mail Questionnaire, Socioly and Social Research, Vol. 50, Oct., 1965.

14- Rummel, J. and Ballaine, W., Research Methodology in Business, New York : Harper & Row, 1963.

15- Schuessler, K., Analyzing Social Data, Boston : Houghton Mifflin, 1971.

16- Selltize, C. et al., Research Methods in Social Relations, New York : Holt, 1959.

17- Spear, M., Practical Charting Techniques, New York : McGraw-Hill. 1969.

18- Spiegel, M., Theory and Problems of Statistics, New York : Schaum Publishing Co., 1961.

19- Stockton, J., Business Statistics, 2nd ed. Cincinati, Ohio : South Western Publishing Co., 1962.

20- Stover, R. and Stone, W., Hand Delivery of Self-Administered Questionnaires, Public Opinion Quarterly, Vol. 38, Summer, 1974.

21- Takeuchi, H. and Schmidt, A., New Promise of Computer Graphics, Harvard Business Review, Jan. - Feb., 1980.

22- Tyebjee, T., Telephone Survey Method : The State of the Art, Journal of Marketing, Vol. 43, Summer, 1979.

تصميم مؤسسة الشراع للخدمات المطبعية
جبل الحسين-تلفاكس ٥ ٦٩٣٤٠-ص.ب ٤٢٦١٠٨ عمان ١١١٤٠ الأردن

SCIENTIFIC RESEARCH METHODS
For
ADMINISTRATIVE SCIENCES

Dr.
ALI SALEEM AL-ALAWNEH
University of Muta

First Edition 1996

Dar Al-Fikr
AMMAN - JORDAN